

DUALITÄT UND APPROXIMATION
BEI
PARABOLISCHEN KONTROLLPROBLEMEN

Uwe Mackenroth

Mainz 1976

DUALITÄT UND APPROXIMATION
BEI
PARABOLISCHEN KONTROLLPROBLEMEN

D i s s e r t a t i o n

zur Erlangung des Grades

"D o k t o r d e r N a t u r w i s s e n s c h a f t"

am Fachbereich Mathematik

der Johannes Gutenberg Universität in Mainz

Uwe Mackenroth
geb. in Hamburg

Mainz 1976



Berichterstatter :

Prof. Dr. H.N. Mülthei, Mainz

Prof. Dr. W. Krabs, Darmstadt

Prof. Dr. F. Lempio, Würzburg

Datum der mündlichen Prüfung : 14.7.1976

INHALTSVERZEICHNIS

EINLEITUNG	1
KAPITEL I. DUALITÄT BEI PARABOLISCHEN KONTROLLPROBLEMEN	
§1. Lineare Differentialgleichungen in Sobolev-Räumen	7
1.1. Sobolev-Räume	7
1.2. Vektorwertige Distributionen	9
1.3. Elliptische Differentialgleichungen	10
1.4. Parabolische Differentialgleichungen	13
§2. Parabolische Kontrollprobleme	15
2.1. Das allgemeine Problem	15
2.2. Ein Transmissions-Kontrollproblem	15
§3. Konvexe Optimierung und Dualität	20
§4. Dualität bei parabolischen Kontrollproblemen	23
4.1. Die Herleitung des Dualproblems	23
4.2. Der Dualitätssatz	26
4.3. Beispiele	27
§5. Charakterisierungen optimaler Steuerungen	30
5.1. Das allgemeine Kontrollproblem	30
5.2. Beispiele	31
KAPITEL II. APPROXIMATION PARABOLISCHER KONTROLLPROBLEME	
§6. Approximation konvexer Optimierungsprobleme	33
6.1. Konvergenzsätze	33
6.2. Der Beweis von Satz 6.1	35
6.3. Ein Hilfssatz	38
§7. Diskretisierungen parabolischer Kontrollprobleme	40
7.1. Das diskrete primale Problem	40
7.2. Das Dualproblem des diskretisierten Kontrollproblems	41
§8. Approximation der Menge der zulässigen Steuerungen	45
8.1. Allgemeine Voraussetzungen	45
8.2. Beispiele	46
§9. Galerkin-Approximation parabolischer Differentialgleichungen	51

§10. Konvergenzaussagen für das parabolische Kontrollproblem	57
10.1. Konvergenz der Extremalwerte	57
10.2. Konvergenz der Optimallösungen	59
§11. Ein numerisches Beispiel	62
11.1. Das diskrete Problem in Matrizenschreibweise	62
11.2. Die numerischen Ergebnisse	64
LITERATURVERZEICHNIS	68

EINLEITUNG

In dieser Arbeit betrachten wir ein Kontrollproblem, in dessen Nebenbedingungen eine lineare parabolische Differentialgleichung vorkommt. Die Menge der zulässigen Steuerungen sei ebenso wie die Zielfunktion konvex. Dadurch können wir im ersten Teil der Arbeit die konvexe Optimierungstheorie anwenden, um mit Hilfe eines Dualproblems Charakterisierungs- und Existenzsätze zu gewinnen. Im zweiten Kapitel beschäftigen wir uns mit Konvergenzfragen, die bei der Diskretisierung des Problems entstehen.

In jeder Hinsicht fundamental für die Untersuchungen ist der Lösungsbegriff, der der partiellen Differentialgleichung zugrunde liegt. Es sind bei dessen Auswahl eine Reihe von Gesichtspunkten zu berücksichtigen. Auf der einen Seite mag es wünschenswert erscheinen, einen möglichst klassischen Lösungsbegriff zu wählen, da dies bei der Approximation der Lösung im allgemeinen von Vorteil ist. Die meisten Konvergenzsätze verlangen daher auch mehr oder weniger starke Regularität der Lösung. Auf der anderen Seite müssen wir garantieren, daß für jede Steuerung aus einem gewissen Funktionenraum eine Lösung der Differentialgleichung existiert. Es ist bekannt, daß eine optimale Steuerung durchaus unstetig sein kann (z.B., wenn es sich um eine Bang-Bang-Steuerung handelt). Man wird daher den linearen Raum der Steuerungen zumindest so groß wählen, daß er die wesentlich beschränkten Funktionen umfaßt. Dann wird man aber im allgemeinen nicht umhin können, für die Differentialgleichung in geeigneter Weise verallgemeinerte Lösungen zuzulassen.

Es soll darauf hingewiesen werden, daß man unter Umständen den Lösungsbegriff für die Differentialgleichung weitgehend überspielen kann; siehe etwa GLASHOFF, WECK [14] oder auch KRABS, WECK [23], KRABS [22], GLASHOFF, GUSTAFSON [11], GLASHOFF, KRABS [13]. Dort wird zunächst für unendlich oft differenzierbare Randdaten mittels einer Reihenentwicklung ein Lösungsoperator definiert, der sich dann auf den Raum der wesentlich beschränkten Funktionen fortsetzen läßt. Für die Herleitung von Optimalitätskriterien spielt es dann keine Rolle, in welcher Weise man die Werte des Lösungsoperators als Lösung der Differentialgleichung interpretieren kann. Läßt man dagegen in dem von GLASHOFF, WECK [14] behandelten Problem sogar quadratintegrierbare Steuerungen zu, so ergeben sich als Werte für den Lösungsoperator Distributionen (siehe LIONS [25], S.202), was physikalisch natürlich wenig sinnvoll ist. Bei diesem Problem kommt man mit dem Reihenansatz weiter als mit der in LIONS - MAGENES [31], [32] darge-

stellten allgemeinen Theorie der partiellen Differentialgleichungen. Dort müssen die Daten aus gewissen Hilbert-Räumen stammen, so daß für die Steuerungen in den meisten Fällen nur quadratintegrierbare Funktionen in Frage kommen.

Andererseits gibt es eine Fülle von Kontrollproblemen, für die auch quadratintegrierbare Steuerungen sinnvoll sind. Beispiele dafür findet man in LIONS [25]. Auch unser in § 2.2 eingeführtes Modellproblem fällt darunter. Weiter ist ein wesentliches Merkmal der in LIONS - MAGENES [31], [32] entwickelten Theorie, daß zu der gegebenen Differentialgleichung eine äquivalente Variationsform aufgestellt wird. Damit liegt die Approximation der Lösung mittels eines Galerkin-Ansatzes auf der Hand. Dagegen wird die numerische Auswertung der Fourierreihe nur in wenigen günstig gelagerten Fällen ein geeigneter Weg zur Approximation der Lösung sein.

Aus diesen Gründen werden wir uns auf die Theorie aus LIONS - MAGENES [31], [32] stützen und verzichten auch nicht darauf, die für uns wichtigsten Ergebnisse in einem einleitenden Abschnitt zusammenzufassen. Unter anderem soll dabei die Wahl der verwendeten Funktionenräume begründet werden. Auf die Darstellung tiefer liegender Ergebnisse, wie etwa der Regularitätssätze oder der Green'schen Formel in Sobolev-Räumen verzichten wir, da sich diese auf unser Modellproblem zumindest nicht unmittelbar anwenden lassen. Die Ursache dafür ist, daß die Koeffizienten der Differentialgleichung aus dem Modellproblem nur stückweise unendlich oft differenzierbar sind, die genannten Ergebnisse diese Eigenschaft aber in dem gesamten Gebiet fordern.

Im zweiten Paragraphen sind wir dann in der Lage, das Kontrollproblem zu formulieren. Unser Problem hat dabei zwangsläufig eine ähnliche Struktur wie das in LIONS [25], Kapitel III, eingeführte, da die Wahl der Funktionenräume im wesentlichen durch die Differentialgleichung festgelegt wird und wir Restriktionen für die Zustandsvariable nicht zulassen wollen. Im Gegensatz zu LIONS [25] betrachten wir aber auch Fälle, in denen nicht das Quadrat einer Norm minimiert wird. So lassen wir z. B. auch die Norm selbst zu. Die Zielfunktion ist dann im allgemeinen auf einem Banach-Raum definiert.

Da unser Problem konvex ist, kann man ein duales Problem aufstellen. Mit Hilfe dieses Dualproblems lassen sich Charakterisierungs- und Existenzaussagen für das primale Problem gewinnen und man kann untere Schranken für das Infimum des primalen Problems angeben. Insbesondere erweist sich das Dualproblem bei der Untersuchung von Approximationsfragen als sehr nützlich. Daher ist unser erstes Ziel die Aufstellung eines Dualproblem. Vorher geben

wir eine kurze Einführung in die Dualitätstheorie, die sich im wesentlichen an ROCKAFELLAR [35] anschließt.

Von den Autoren, die sich mit Dualität bei Kontrollproblemen mit partiellen Differentialgleichungen beschäftigen, nennen wir zuerst LIONS [25]. Das dort hergeleitete Dualproblem hat aber gegenüber dem unsrigen den Nachteil, daß es für die Berechnung der dualen Zielfunktion die Kenntnis der Optimallösung des bezüglich der Steuervariablen unrestringierten Primalproblems erfordert. Außerdem ist es nur auf den Fall einer strikt konvexen Zielfunktion anwendbar. Ein weiteres Dualproblem findet sich in LIONS [27], wo eine von ROCKAFELLAR [34] gegebene Verallgemeinerung des Fenchel'schen Dualproblems auf gewisse elliptische Kontrollprobleme angewendet wird. Für ein von dem unsrigen abweichendes parabolisches Kontrollproblem wird in KRABS [22], GLASHOFF, KRABS [13] ein Dualproblem aufgestellt. Zu dem in §4 erhaltenen Ergebnis besteht aber insofern eine gewisse Ähnlichkeit, als daß auch dort im Falle der Minimumnormprobleme über die Einheitskugel eines gewissen Dualraumes zu maximieren ist (vergl. §4.3, Beispiel 2).

Mit Hilfe des Dualproblems werden in §5 notwendige und hinreichende Bedingungen für Optimalität hergeleitet. Im Spezialfall der quadratischen Probleme ergeben sich dabei die in LIONS [25] mit Differenzierbarkeitsargumenten bewiesenen Charakterisierungen. Da unser Vorgehen aber auf keinerlei Differenzierbarkeitsvoraussetzungen basiert, sind wir auch in der Lage, gewisse nicht-differenzierbare Zielfunktionen zu behandeln. Als Beispiel betrachten wir in §5.2 ein Minimumnormproblem. Insbesondere erhalten wir in §5 für unser Modellproblem eine Charakterisierung. Die hieraus gewonnene Information über eine optimale Steuerung verwenden wir später bei der Diskretisierung des Modellproblems.

Im zweiten Kapitel führen wir die Diskretisierung für das Kontrollproblem ein und untersuchen das Problem, wann die diskreten Infima gegen das kontinuierliche streben. Wir greifen diese Frage zunächst für allgemeine konvexe Optimierungsprobleme auf. Es gibt bereits eine Reihe von Autoren, die sich in sehr allgemeiner Weise auch mit Konvergenzuntersuchungen bei nicht-konvexen Problemen beschäftigen, so etwa DANIEL [8], KRABS [21]. Die speziellen Eigenschaften konvexer Probleme werden dagegen meist nicht beachtet. KRABS [19] beschränkt sich zwar auf konvexe Probleme, aber eines der dortigen Hauptergebnisse (Satz 3.1) wird in KRABS [21] mit im wesentlichen denselben Voraussetzun-

gen auf nicht-konvexe Probleme verallgemeinert. Die Ausnutzung des Dualproblems beim Beweis von Satz 2.3 in KRABS [19] ist offenbar mehr beweistechnischer Natur. Es zeigt sich aber, daß man die Konvexität in sehr viel wesentlicherer Weise beim Beweis eines Konvergenzsatzes anwenden kann. So verlangen wir nur einen Teil der in KRABS [21] zugrunde gelegten Hausdorff-Konvergenz der diskreten Restriktionenmengen C_n gegen die kontinuierliche Restriktionenmenge C . Dafür müssen wir die analoge Voraussetzung für die Restriktionenmengen im Dualproblem stellen. Welche der Annahmen schwächer ist, die der vollen Hausdorff-Konvergenz oder unsere Voraussetzungen (V1), (V1_{*}) aus §6, kann an dieser Stelle nicht entschieden werden, aber die folgende Anwendung auf das Kontrollproblem zeigt, daß der Weg über das Dualproblem durchaus praktikabel ist. Dagegen benötigen wir keinerlei Voraussetzungen über die gleichmäßige Konvergenz der diskreten Zielfunktionen gegen die kontinuierliche, wie etwa KRABS [21] oder implizit auch DANIEL [8]. Der Grund dafür ist, daß man diese Voraussetzung abschwächen kann und eine entsprechende Annahme für das Dualproblem macht. Die abgeschwächte Voraussetzung können wir dann generell für konvexe Funktionale aus punktwiser Konvergenz folgern.

In §7 diskretisieren wir das Kontrollproblem und stellen die Dualprobleme der diskreten Probleme auf. Für die Menge der zulässigen Steuerungen lassen wir relativ allgemeine Approximationen zu. Die genauen Voraussetzungen hierfür formulieren wir in §8 und betrachten dort auch einige Beispiele. Wir verzichten aber auf eine Diskretisierung der Zielfunktion, da man diese in vielen praktischen Fällen exakt auswerten kann. Problematischer dagegen ist die Diskretisierung der Differentialgleichung, da wir keinerlei Regularitätsaussagen verwenden. Wir haben bereits erwähnt, daß ein Galerkin-Ansatz mit Finiten Elementen nahe liegt. Bei den in der Literatur behandelten Problemen dieser Art geht man dabei meist so vor, daß die Zeitvariable zunächst nicht diskretisiert wird, und man einen Finite-Elemente-Ansatz in den Ortsvariablen macht (vergl. DOUGLAS, DUPONT [9]). Für das so entstehende System gewöhnlicher Differentialgleichungen wird dann ein Differenzenverfahren angesetzt. Andererseits ist es ebenfalls sinnvoll, einen Galerkin-Ansatz sowohl für die Orts- als auch für die Zeitvariable zu machen. Zumindest unter den hier betrachteten Voraussetzungen für die Differentialgleichung scheint dieser Weg noch nicht beschritten worden zu sein, so daß wir in §9 einen Konvergenzbeweis geben.

In §10 sind wir schließlich in der Lage, die Konvergenz der diskreten Minima gegen das kontinuierliche für das Kontrollproblem zu beweisen. Wir erhalten auch einige Ergebnisse, die die Konvergenz der Optimallösungen betreffen. Schwache Konvergenz der diskreten Optimallösungen gegen eine kontinuierliche ergibt sich ohne Mühe, der Schluß auf Normkonvergenz ist dagegen problematischer. Relativ leicht erhält man sie, wenn die Zielfunktion koerzitiv ist (vergl. Voraussetzung 2.2). Im nicht-koerzitiven Fall erhält man Normkonvergenz im wesentlichen aus dem Bang-Bang-Prinzip, falls dies erfüllt ist.

GLASHOFF und KRABS [12] beschäftigen sich ebenfalls mit der Diskretisierung von parabolischen Kontrollproblemen. Bei dem dort gegebenen Problem geht die Steuerung nicht direkt in die Daten der Differentialgleichung ein, sondern wird erst durch einen Operator geglättet. Infolgedessen erhält man für die Differentialgleichung klassische Lösungen, die mit einem gewöhnlichen Differenzenverfahren approximiert werden. Daher sind diese Ergebnisse nicht direkt mit den unsrigen vergleichbar, ebensowenig wie die von GLASHOFF und GUSTAFSON [11], wo man die Lösung der Differentialgleichung exakt kennt und als diskrete zulässige Steuerungen nur solche verwendet, die auch für das kontinuierliche Problem zulässig sind.

BOSARGE, JOHNSON, SMITH [4] diskretisieren Steuerungsprobleme, bei denen die Steuerungen keiner Restriktion unterworfen werden und die Zielfunktion als koerzitiv vorausgesetzt ist. Die Anwendbarkeit dieser Ergebnisse auf praktische Probleme wird aber durch derartige Voraussetzungen stark begrenzt. Insbesondere werden so die Fälle ausgeschlossen, in denen das Bang-Bang-Prinzip gilt. Es ist auch nicht zu vermuten, daß sich die Vorgehensweise von BOSARGE, JOHNSON, SMITH [4] auf derartige Fälle ausdehnen läßt, da die Koerzitivität der Zielfunktion in wesentlicher Weise in den Konvergenzbeweis zur Gewinnung von Fehlerabschätzungen eingeht.

Die von BENSOUSSAN, BOSSAVIT, NEDELEC [3] betrachtete Problemstellung kommt der unsrigen am nächsten. Im koerzitiven Fall kommen die Autoren im wesentlichen zu demselben Ergebnis wie wir, allerdings auf ganz anderem Wege. Sie beweisen nicht erst die Konvergenz der diskreten Minima gegen das kontinuierliche, sondern zeigen direkt die Konvergenz der Optimallösungen. Da sich diese Methode nicht auf den nicht-koerzitiven Fall verallgemeinern läßt, "regularisieren" die Autoren das Problem, indem sie einen von einer positi-

ven Zahl λ abhängigen Term zur alten Zielfunktion addieren, so daß das neue Problem koerzitiv wird. Für $\lambda = 0$ ergibt sich wieder das alte Problem. Die Autoren zeigen, daß die Optimallösungen des diskreten regularisierten Problems gegen eine Optimallösung des Ausgangsproblems streben, wenn λ in geeigneter Weise gegen Null strebt. Diese Regularisierung ist aber insofern fragwürdig, als daß sie nur beweistechnisch begründet ist, nicht aber durch das Ziel, das infinite Problem durch ein finites Problem zu ersetzen. Außerdem hat man bei der numerischen Behandlung einen Diskretisierungsparameter mehr zu berücksichtigen.

Im letzten Abschnitt wird das in §2 beschriebene Transmissions-Kontrollproblem im Sinne von §7 diskretisiert und mit einem Gradientenverfahren numerisch gelöst. Dabei zeigt es sich, daß die numerischen Ergebnisse die Konvergenzsätze 10.1 und 10.5, die die Konvergenz der Minimalwerte betreffen, bestätigen.

An dieser Stelle möchte ich Herrn Professor Dr. Mülthei herzlich für die Betreuung dieser Arbeit und eine Reihe sehr nützlicher Gespräche danken. Mein besonderer Dank gilt auch Herrn Professor Dr. Krabs, der mir eine Reihe wertvoller Anregungen gegeben hat.

I. DUALITÄT BEI PARABOLISCHEN KONTROLLPROBLEMEN

§1. Lineare Differentialgleichungen in Sobolev-Räumen

1.1. Sobolev-Räume

Sei Ω eine offene Teilmenge des $\mathbb{R}^n$ mit Rand Γ .

Mit $\mathcal{D}(\Omega)$ bezeichnen wir den Raum der unendlich oft differenzierbaren reellwertigen Funktionen auf Ω mit kompaktem Träger in Ω . Topologisiert man $\mathcal{D}(\Omega)$ in der bekannten Weise, so ist der Dualraum $\mathcal{D}'(\Omega)$ von $\mathcal{D}(\Omega)$ der Raum der Distributionen auf Ω (vergl. SCHWARTZ [36]).

$L^2(\Omega)$ sei der Hilbert-Raum (der Äquivalenzklassen) der meßbaren quadratintegrierbaren Funktionen auf Ω mit dem üblichen Skalarprodukt. $L^2(\Omega)$ läßt sich bekanntlich als Teilraum von $\mathcal{D}'(\Omega)$ auffassen, so daß man die Inklusionen

$$(1.1) \quad \mathcal{D}(\Omega) \subset L^2(\Omega) \subset \mathcal{D}'(\Omega)$$

hat und daher distributionelle Ableitungen $\frac{\partial u}{\partial x_i}$ von Funktionen $u \in L^2(\Omega)$ bilden kann.

Für $\alpha \in (\mathbb{N} \cup \{0\})^n$ setzen wir

$$|\alpha| := \sum_{i=1}^n \alpha_i,$$

$$D^\alpha := D_1^{\alpha_1} \dots D_n^{\alpha_n}, \quad D_i := \frac{\partial}{\partial x_i},$$

und definieren für $m \in \mathbb{N} \cup \{0\}$

$$(1.2) \quad H^m(\Omega) := \{u \in L^2(\Omega) \mid D^\alpha u \in L^2(\Omega) \forall \alpha, |\alpha| \leq m\}.$$

$H^m(\Omega)$ heißt Sobolev-Raum der Ordnung m auf Ω . Mit dem inneren Produkt

$$(1.3) \quad \langle u, v \rangle_{H^m(\Omega)} := \sum_{|\alpha| \leq m} \langle D^\alpha u, D^\alpha v \rangle_{L^2(\Omega)}$$

wird er zum Hilbert-Raum.

Mit $H_0^m(\Omega)$ bezeichnen wir den Abschluß von $\mathcal{D}(\Omega)$ in $H^m(\Omega)$ und mit $H^{-m}(\Omega)$ den Dualraum von $H_0^m(\Omega)$. Man kann zeigen, daß $H^{-m}(\Omega)$ mit dem Teilraum derjenigen Distributionen $T \in \mathcal{D}'(\Omega)$ identifiziert werden kann, die sich in der Form

$$(1.4) \quad T = \sum_{|\alpha| \leq m} D^\alpha f_\alpha, \quad f_\alpha \in L^2(\Omega),$$

darstellen lassen. Für $u \in H_0^m(\Omega)$ ist $T(u)$ dabei folgendermaßen zu verstehen. Da $\mathcal{D}(\Omega)$ dicht in $H_0^m(\Omega)$ liegt, gibt es Elemente $\varphi_i \in \mathcal{D}(\Omega)$, die in der Topologie von $H^m(\Omega)$ gegen u konvergieren. Wie sich zeigen läßt, ist T im Sinne dieser Topologie auf $\mathcal{D}(\Omega)$ stetig, so daß man

$$T(u) := \lim_{i \rightarrow \infty} T(\varphi_i)$$

setzen kann.

Wir machen jetzt für Ω die folgenden Voraussetzungen:

(1.5) Ω ist beschränkt;

der Rand Γ von Ω ist eine $(n-1)$ -dimensionale C^∞ -Mannigfaltigkeit,

(1.6) so daß Ω lokal auf einer Seite von Γ liegt.

Dann wird mit $\mathcal{D}(\Gamma)$ der Raum der unendlich oft differenzierbaren Funktionen auf Γ bezeichnet. Schließlich benötigen wir noch den Raum $\mathcal{D}(\bar{\Omega})$ aller Funktionen u auf Ω , zu denen es ein $v \in \mathcal{D}(\mathbb{R}^n)$ mit $u = v|_\Omega$ gibt. Für $u \in \mathcal{D}(\bar{\Omega})$ setze

$$(1.7) \quad \gamma(u) := u|_\Gamma.$$

Mit diesen Vorbereitungen können wir nun einen Spezialfall des für die Theorie der partiellen Differentialgleichungen fundamentalen Spursatzes zitieren (LIONS - MAGENES [31], Satz 8.3).

Satz 1.1. Erfülle Ω die Voraussetzungen (1.5), (1.6). Dann läßt sich die durch (1.7) definierte Abbildung $\gamma : \mathcal{D}(\bar{\Omega}) \rightarrow \mathcal{D}(\Gamma)$ in eindeutiger Weise zu einer stetigen Abbildung von $H^m(\Omega)$ nach $L^2(\Gamma)$ fortsetzen ($m \in \mathbb{N}$). Der Kern von γ ist $H_0^m(\Omega)$.

Hierbei bezeichnet $L^2(\Gamma)$ den Raum (der Äquivalenzklassen) der bezüglich des vom Lebesgue-Maß dx induzierten Oberflächenmaßes $d\Gamma$ meßbaren und quadratintegrierbaren Funktionen auf Γ .

Schließlich definieren wir $L^\infty(\Omega)$ als den Banach-Raum (der Äquivalenzklassen) der meßbaren und wesentlich beschränkten Funktionen auf Ω mit der üblichen Norm.

1.2. Vektorwertige Distributionen

Seien A und B topologische lineare Räume. Den Raum der stetigen linearen Operatoren von A nach B bezeichnen wir mit $\mathcal{L}(A, B)$.

Sei Z ein Hilbert-Raum. Dann wird für reelles $T > 0$ der Raum der Distributionen auf $]0, T[$ mit Werten in Z durch

$$(1.8) \quad \mathcal{D}^*(]0, T[, Z) := \mathcal{L}(\mathcal{D}([0, T]), Z)$$

definiert. Für $S \in \mathcal{D}^*(]0, T[, Z)$ läßt sich analog zum reellwertigen Fall die distributionelle Ableitung $\frac{dS}{dt}$ von S durch

$$(1.9) \quad \frac{dS}{dt}(\varphi) := -S\left(\frac{d\varphi}{dt}\right) \quad \forall \varphi \in \mathcal{D}([0, T])$$

erklären (vergl. SCHWARTZ [37]).

Ein wichtiger Teilraum von $\mathcal{D}^*(]0, T[, Z)$ ist der Hilbert-Raum $L^2(0, T; Z)$ (der Äquivalenzklassen) der stark meßbaren Funktionen $u:]0, T[\rightarrow Z$ mit

$$(1.10) \quad \int_0^T \|u(t)\|_Z^2 dt < \infty.$$

Die Integralbildung in (1.10) ist sinnvoll, da $t \mapsto \|u(t)\|_Z^2$ meßbar ist. Überdies existiert für alle $u \in L^2(0, T; Z)$ das Bochner-Integral

$$\int_0^T u(t) dt,$$

das ein Element aus Z ist (YOSHIDA [40], Kap. V). Daher kann man zu jedem $u \in L^2(0, T; Z)$ eine vektorwertige Distribution

$$(1.11) \quad S_u(\varphi) := \int_0^T u(t) \varphi(t) dt$$

konstruieren und erhält so eine stetige, injektive Einbettung

$$(1.12) \quad L^2(0, T; Z) \subset \mathcal{D}^*(]0, T[, Z).$$

Seien Hilbert-Räume X, Y mit $X \subset Y$ gegeben. X sei dicht in Y und die Inklusion sei stetig. Dann setzen wir

$$(1.13) \quad W(0, T; X, Y) := \{u \in L^2(0, T; X) \mid \frac{du}{dt} \in L^2(0, T; Y)\}.$$

In $W(0, T; X, Y)$ liegen also alle $u \in L^2(0, T; X)$, zu denen ein $g \in L^2(0, T; Y)$ mit

$$-\int_0^T u(t) \dot{\varphi}(t) dt = \int_0^T g(t) \varphi(t) dt \quad \forall \varphi \in \mathcal{D}([0, T])$$

existiert, wobei die Integrale im Sinne von $L^2(0,T;Y)$ zu nehmen sind. Mit dem Skalarprodukt

$$(1.14) \quad \langle u, v \rangle_{W(0,T;X,Y)} := \int_0^T \langle u(t), v(t) \rangle_X dt + \int_0^T \langle \frac{du}{dt}(t), \frac{dv}{dt}(t) \rangle_Y dt$$

wird $W(0,T;X,Y)$ zum Hilbert-Raum.

Beispiel: Sei $Q :=]0, T[\times \Omega$. Dann wird durch

$$(1.15) \quad j(u)(t) := u(t, \cdot), \quad t \in]0, T[,$$

ein isometrischer Isomorphismus j von $H^1(Q)$ auf $W(0,T;H^1(\Omega), L^2(\Omega))$ definiert. Dabei gilt

$$(1.16) \quad \left(\int_0^T j(u)(t) dt \right)(x) = \int_0^T u(t, x) dt \quad \forall x \in \Omega,$$

unabhängig, ob das Bochner-Integral in $L^2(0,T;H^1(\Omega))$ oder $L^2(0,T;L^2(\Omega))$ genommen wird. Für die vektorielle und die gewöhnliche distributionelle Ableitung gilt die Beziehung

$$(1.17) \quad \frac{dj(u)}{dt}(t) = \frac{\partial u}{\partial t}(t, \cdot).$$

Man kann zeigen, daß zu jedem $u \in W(0,T;X,Y)$ ein $a \in Y$ existiert, so daß

$$(1.18) \quad u(t) = a + \int_0^t \frac{du}{d\tau}(\tau) d\tau \quad \text{f.ü. auf }]0, T[.$$

Die rechte Seite ist stetig in t auf $[0, T]$. Daher kann u nach einer eventuellen Abänderung auf einer Nullmenge als stetige Funktion von $[0, T]$ nach Y angesehen werden. Insbesondere ist also $u(0)$ definiert und es gilt $u(0) \in Y$.

1.3. Elliptische Differentialgleichungen

In diesem Abschnitt soll, ausgehend von einem Beispiel, ein für die Behandlung elliptischer Differentialgleichungen geeigneter funktionalanalytischer Rahmen angegeben werden. Dieser wird dann in § 1.4 auf den uns eigentlich interessierenden parabolischen Fall ausgedehnt.

Sei $C_0^2(\bar{\Omega})$ die Menge der auf $\bar{\Omega}$ zweimal stetig differenzierbaren Funktionen, die auf Γ verschwinden. Versieht man $C_0^2(\bar{\Omega})$ mit der üblichen Maximumnorm, so

ist $C_0^2(\bar{\Omega})$ ein Banach-Raum. Für $e \in C(\bar{\Omega})$ und $c > 0$ sei die Randwertaufgabe

$$(1.19) \quad -\Delta u + cu = e, \quad u \in C_0^2(\bar{\Omega}),$$

gegeben. Um zu zeigen, daß (1.19) eine Lösung hat, kann man diese Differentialgleichung in ein Variationsproblem umformulieren und dessen Lösbarkeit beweisen. Das zu (1.19) äquivalente Variationsproblem besteht bekanntlich darin, das Funktional

$$(1.20) \quad a(u, u) - 2 \langle u, e \rangle_{L^2(\Omega)}$$

über alle $u \in C_0^2(\bar{\Omega})$ zu minimieren, wobei

$$(1.21) \quad a(u, v) := \sum_{i=1}^n \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i} \right\rangle_{L^2(\Omega)} + c \langle u, v \rangle_{L^2(\Omega)}$$

sei. Um die Lösbarkeit dieser Variationsaufgabe zu beweisen, könnte man wie folgt vorgehen. Man wählt eine Minimalfolge und überprüft, ob diese im Sinne einer geeigneten Topologie von $C_0^2(\bar{\Omega})$ in einer folgenkompakten Menge von $C_0^2(\bar{\Omega})$ liegt. Wenn dies der Fall ist, existiert eine gegen ein gewisses $u \in C_0^2(\bar{\Omega})$ konvergente Teilfolge der Minimalfolge und man kann vermuten, daß u Minimal-lösung ist.

Dieser Kompaktheitsschluß läßt sich in $C_0^2(\bar{\Omega})$ offenbar nicht durchführen, wohl aber in $H_0^1(\Omega)$. Man zeigt nämlich leicht, daß ein $\alpha > 0$ mit

$$(1.22) \quad a(u, u) \geq \alpha \|u\|_{H^1(\Omega)}^2 \quad \forall u \in H^1(\Omega)$$

existiert und folgert hieraus, daß die Minimalfolge in einer beschränkten (und abgeschlossenen) Teilmenge von $H_0^1(\Omega)$ liegt. Diese ist schwach folgenkompakt, da $H^1(\Omega)$ Hilbert-Raum ist.

Man kann den angedeuteten Schluß tatsächlich zu Ende führen und so zeigen, daß es ein Element aus $H_0^1(\Omega)$ gibt, das dem jetzt auf $H_0^1(\Omega)$ definierten Funktional (1.20) seinen minimalen Wert erteilt. Dies genügt dann auch der Differentialgleichung (1.19) (im distributionellen Sinn).

Auf diese Weise bietet sich also $H_0^1(\Omega)$ in natürlicher Weise als Lösungsraum der Differentialgleichung (1.19) an. Weiter spielt es keine Rolle mehr, daß e in $C(\bar{\Omega})$ liegt. Das Funktional (1.20) hat auch für $e \in H_0^1(\Omega)'' = H^{-1}(\Omega)$

einen Sinn, ebenso die Differentialgleichung. Dann existiert ebenfalls eine Lösung in $H_0^1(\Omega)$.

Um auch größere Klassen von elliptischen Differentialgleichungen zu erfassen, ist es erforderlich, die obige Vorgehensweise in einem allgemeinen funktional-analytischen Rahmen zu formulieren.

Seien Hilbert-Räume V und H gegeben. Für die Skalarprodukte und Normen auf V und H verwenden wir die folgenden Bezeichnungen:

$$(1.23) \quad \langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_V, \quad (\cdot, \cdot) := \langle \cdot, \cdot \rangle_H, \quad \|\cdot\| := \|\cdot\|_V, \quad |\cdot| := \|\cdot\|_H.$$

Weiter gelte

$$(1.24) \quad \begin{aligned} &V \subset H; \quad V \text{ liegt dicht in } H; \\ &\text{die Injektion von } V \text{ in } H \text{ ist stetig.} \end{aligned}$$

Sei $h \in H$. Dann ist $(\cdot, h)$ ein stetiges lineares Funktional auf V . Identifiziert man H^* mit H , so gilt also

$$(1.25) \quad V \subset H \subset V^*,$$

wobei die Inklusionen stetig sind und jeder Raum in dem folgenden dicht liegt. Wir bezeichnen auch die Dualität von V und V^* mit $(\cdot, \cdot)$. Dies ist mit der vorigen Definition von $(\cdot, \cdot)$ konsistent, da für $h \in H$ die Gleichung $(v, h) = (v, h)_H$ für alle $v \in V$ gilt. Sei a eine stetige Bilinearform auf $V \times V$, die koerzitiv ist, d.h. für die ein $\alpha > 0$ mit

$$(1.26) \quad a(v, v) \geq \alpha \|v\|^2 \quad \forall v \in V$$

existiert. Dann gilt der

Satz 1.2. Unter den obigen Voraussetzungen gibt es zu jedem $e \in V^*$ genau ein $u \in V$ mit

$$(1.27) \quad a(v, u) = (v, e) \quad \forall v \in V.$$

Falls a symmetrisch ist, gilt (1.27) genau dann, wenn u dem Funktional $v \mapsto a(v, v) - 2(v, e)$ den kleinsten Wert erteilt und man kann in diesem Fall Satz 1.2 leicht mit Hilfe des Variationsproblems beweisen.

Für festes $u \in V$ ist durch $a(\cdot, u)$ ein Element aus V^* gegeben. Man kann daher einen Operator $A \in \mathcal{L}(V, V^*)$ durch $Au := a(\cdot, u)$ definieren. Dann gilt also

$$(1.28) \quad a(v, u) = (v, Au) \quad \forall u, v \in V$$

und (1.28) ist mit der Operatorgleichung

$$(1.29) \quad Au = e$$

äquivalent.

Das Beispiel zu Beginn fügt sich in diese Theorie ein, wenn man $V = H_0^1(\Omega)$ und $H = L^2(\Omega)$ wählt. Bei allgemeineren elliptischen Differentialgleichungen zweiter Ordnung ist V ein abgeschlossener Teilraum von $H^1(\Omega)$, der $H_0^1(\Omega)$ umfaßt. Man beachte, daß A dann im allgemeinen nicht mit dem Differentialoperator der Randwertaufgabe identisch ist, weil in A noch Randterme eingehen. An diesem Beispiel sieht man auch, warum V^* nicht mit V identifiziert wurde. Man hätte sonst nämlich einen Operator $\mathcal{O}: V \rightarrow V$ mit $a(v, u) = \langle v, \mathcal{O}u \rangle$ erhalten, der von dem Differentialoperator aus (1.19) verschieden gewesen wäre.

1.4. Parabolische Differentialgleichungen

Wir beginnen wieder mit einem Beispiel. Gesucht ist eine Funktion y mit

$$\frac{\partial y}{\partial t}(t, x) - \Delta y(t, x) + cy(t, x) = e(t, x) \quad \forall (t, x) \in]0, T[\times \Omega,$$

$$(1.30) \quad \begin{aligned} y(t, x) &= 0 & \forall (t, x) \in]0, T[\times \Gamma, \\ y(0, x) &= 0 & \forall x \in \Omega. \end{aligned}$$

Man kann jetzt, motiviert durch die Überlegungen aus §1.2, $\bar{y}(t) = y(t, \cdot)$ und $\bar{e}(t) = e(t, \cdot)$ definieren und (1.30) durch

$$\frac{d\bar{y}}{dt}(t) - \Delta \bar{y}(t) + c\bar{y}(t) = \bar{e}(t) \quad \forall t \in]0, T[,$$

$$(1.31) \quad \bar{y}(0) = 0,$$

ersetzen, wobei $\bar{y}$ in $W(0, T; H_0^1(\Omega), H^{-1}(\Omega))$ liegen soll. Dann kann $\bar{y}$ als stetige Funktion angesehen werden, so daß $\bar{y}(0)$ einen Sinn hat (vergl. §1.2). Der Raum $H^{-1}(\Omega)$ tritt auf, weil der Laplace-Operator auf $H_0^1(\Omega)$ seine Werte in $H^{-1}(\Omega)$ hat.

Diese Betrachtungen legen es nahe, parabolische Differentialgleichungen in der folgenden Weise allgemein zu formulieren.

Seien V und H Hilbert-Räume wie im vorigen Abschnitt. Zusätzlich sei V separabel. Zu jedem $t \in]0, T[$ sei eine stetige Bilinearform $a(t, \cdot, \cdot)$ auf $V \times V$ gegeben, so daß gilt:

(1.32) Für alle $u, v \in V$ ist $a(\cdot, u, v)$ meßbar;

es gibt eine Konstante c , so daß

(1.33) $|a(t, u, v)| \leq c \|u\| \|v\| \quad \forall t \in]0, T[, \forall u, v \in V;$

es gibt ein $\lambda \in \mathbb{R}$ und ein $\alpha > 0$, so daß

(1.34) $a(t, u, u) + \lambda |u|^2 \geq \alpha \|u\|^2 \quad \forall t \in]0, T[, \forall u \in V.$

Da für festes t die Bilinearform $a(t, \cdot, \cdot)$ stetig ist, gibt es einen Operator $A(t) \in \mathcal{L}(V, V^*)$ mit

(1.35) $a(t, v, u) = (v, A(t)u) \quad \forall u, v \in V.$

Man zeigt leicht, daß für alle $y \in L^2(0, T; V)$ die Funktion $t \mapsto A(t)y(t)$ in $L^2(0, T; V^*)$ liegt. Sei

(1.36) $W(0, T) := W(0, T; V, V^*) .$

Wir haben schon gesehen, daß die Funktionen y aus $W(0, T)$ nach einer möglichen Abänderung auf einer Nullmenge stetig von $[0, T]$ nach V^* sind. Man kann sogar zeigen, daß sie stetige Funktionen von $[0, T]$ nach H sind (LIONS - MAGENES [31], Kap. I). Dann ist $y(0)$ definiert und liegt in H . Umgekehrt gibt es zu jedem $y_0 \in H$ ein $y \in W(0, T)$ mit $y(0) = y_0$.

Es gilt der folgende

Satz 1.3. Unter den in diesem Abschnitt gemachten Voraussetzungen gibt es zu beliebigem $e \in L^2(0, T; V^*)$ und $y_0 \in H$ genau ein $y \in W(0, T)$, das der Anfangswertaufgabe

(1.37) $\frac{dy}{dt} + A(t)y = e,$

(1.38) $y(0) = y_0,$

genügt.

§2. Parabolische Kontrollprobleme

2.1. Das allgemeine Problem

Seien Hilbert-Räume V, H und Bilinearformen $a(t, \cdot, \cdot)$ wie in § 1.4 gegeben.

Sei $e \in L^2(0, T; V^*)$ und $y_0 \in H$.

Weiter seien ein Banach-Raum $\mathcal{B}$, ein Hilbert-Raum U und stetige lineare Operatoren

$$(2.1) \quad B: U \rightarrow L^2(0, T; V^*), \quad C: L^2(0, T; V) \rightarrow \mathcal{B},$$

$$N: U \rightarrow U$$

gegeben. N sei symmetrisch und positiv semidefinit.

$U_{ad} \subset U$ sei eine nichtleere, abgeschlossene und konvexe Menge. Die Elemente aus U_{ad} heißen zulässige Steuerungen. Weiter machen wir die folgende Voraussetzung.

$$(2.2) \quad U_{ad} \text{ ist beschränkt oder es gibt ein } \nu > 0, \text{ so daß } \langle u, Nu \rangle_U \geq \nu \|u\|_U^2$$

für alle $u \in U$ gilt.

Sei $f: \mathcal{B} \rightarrow \mathbb{R}$ ein konvexes und stetiges Funktional und sei z_d ein Element aus $\mathcal{B}$. Dann können wir das folgende Kontrollproblem formulieren.

(P) Minimiere

$$(2.3) \quad f(Cy - z_d) + \langle u, Nu \rangle_U$$

über alle $y \in L^2(0, T; V)$ und $u \in U_{ad}$ unter den Nebenbedingungen

$$(2.4) \quad \frac{dy}{dt} + A(t)y = e + Bu,$$

$$(2.5) \quad y(0) = y_0.$$

Wir bezeichnen (P) als parabolisches Kontrollproblem, da die Differentialgleichung (2.4) parabolisch ist. Schließlich sei noch angemerkt, daß (2.4) mit

$$(2.6) \quad (v, \frac{dy}{dt}(t)) + a(t, v, y(t)) = (v, e(t) + (Bu)(t)) \quad \forall t \in]0, T[, \forall v \in V$$

äquivalent ist. (2.6) wird Variationsform der Differentialgleichung (2.4) genannt.

2.2. Ein Transmissions-Kontrollproblem

(a) Seien Ω_1, Ω_2 offene Teilmengen des $\mathbb{R}^n$ mit (1.5), (1.6), deren Ränder $\partial\Omega_1, \partial\Omega_2$

einen Durchschnitt S haben, der bezüglich $\partial\Omega_1$ (bzw. $\partial\Omega_2$) einen inneren Punkt hat. Wir setzen $\Omega := \Omega_1 \cup \Omega_2$, $\Gamma_i := \partial\Omega_i \setminus S$ ($i=1,2$).

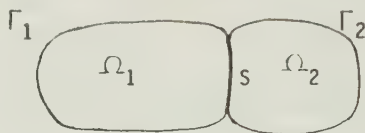


Abb. 1

Wir betrachten das folgende Problem. Das Gebiet $\Omega_1 \cup \Omega_2$ bestehe aus einem Material mit Temperaturleitzahl $k_1(k_2)$ und Wärmeleitzahl $\lambda_1(\lambda_2)$. Das Gebiet $\Omega_1 \cup \Omega_2 \cup S$ sei nach außen vollständig isoliert und in $\Omega_1 \cup \Omega_2$ seien Anfangstemperaturen $y_{01}(y_{02})$ gegeben. Weiter nehmen wir an, daß die in Ω_1 bzw. Ω_2 gegebenen Materialien an der Grenzfläche S in direktem Kontakt miteinander stehen, so daß die Temperatur an der Grenzfläche S stetig verläuft und die in das eine Gebiet einfließende mit der aus dem anderen Gebiet ausfließenden Wärmemenge übereinstimmt. Wird nun in Ω_2 eine gewisse Wärmemenge zugeführt, so stellen sich gemäß den in Ω_1 und Ω_2 gültigen Wärmeleitungsgleichungen und den soeben formulierten Rand-, Anfangs- und Übergangsbedingungen in Ω_1 bzw. Ω_2 Temperaturverteilungen y_1 bzw. y_2 ein. Das Kontrollproblem besteht jetzt darin, eine Wärmezufuhr derart zu bestimmen, daß die sich auf der Grenzfläche S einstellende Temperatur y_1 im quadratischen Mittel möglichst wenig von einer auf S vorgegebenen Temperatur z_d abweicht. Dabei soll die Wärmezufuhr betragsmäßig eine gegebene positive Konstante nicht überschreiten.

Als nächstes geben wir die mathematische Formulierung dieses Problems, wobei wir es für einen Moment noch offen halten, welche Funktionenräume zugrunde liegen.

Minimiere

$$(2.7) \quad \int_0^T \int_S (y_1|_S - z_d)^2 dS dt$$

über alle y_1, y_2, u unter den Nebenbedingungen (für $i = 1, 2$)

$$(2.8) \quad \frac{\partial y_1}{\partial t} = k_1 \Delta y_1,$$

$$(2.9) \quad \frac{\partial y_2}{\partial t} = k_2 \Delta y_2 + u,$$

$$(2.10) \quad \frac{\partial y_i}{\partial n} |_{\Gamma_i} = 0,$$

$$(2.11) \quad y_1|_S = y_2|_S,$$

$$(2.12) \quad \lambda_1 \frac{\partial y_1}{\partial n}|_S + \lambda_2 \frac{\partial y_2}{\partial n}|_S = 0 ,$$

$$(2.13) \quad y_i(0) = y_{0i} ,$$

$$(2.14) \quad |u| \leq m .$$

Hierbei bezeichne $\frac{\partial}{\partial n}$ die Ableitung in Richtung der nach außen orientierten Normalen von Γ . Wir wollen zeigen, daß dieses Problem ein Spezialfall des Problems (P) aus dem vorigen Abschnitt ist. Dazu wird zunächst die folgende Variablentransformation durchgeführt:

$$\tilde{y}_1 := \frac{\sqrt{k_2}}{\sqrt{\lambda_2}} y_1 , \quad \tilde{y}_2 := \frac{\sqrt{k_1}}{\sqrt{\lambda_1}} y_2 .$$

Man erhält wieder die Gleichungen (2.8) - (2.10), (2.13), (2.14), wenn man $\tilde{y}_1, \tilde{y}_2$ in y_1, y_2 umbenennt und u, y_{01}, y_{02} und m in geeigneter Weise umdefiniert. Die Gleichungen (2.11), (2.12) gehen über in

$$(2.11') \quad \sqrt{k_1 \lambda_2} y_1|_S = \sqrt{k_2 \lambda_1} y_2|_S ,$$

$$(2.12') \quad \lambda_1 \sqrt{k_1 \lambda_2} \frac{\partial y_1}{\partial n}|_S + \lambda_2 \sqrt{k_2 \lambda_1} \frac{\partial y_2}{\partial n}|_S = 0 .$$

Als nächstes soll die Variationsform der Differentialgleichungen (2.8) - (2.10), (2.11'), (2.12') aufgestellt werden. Sei $Q_i =]0, T[\times \Omega_i$ und $V_i = H^1(\Omega_i)$, $i = 1, 2$. Für $v \in H^1(\Omega)$ setzen wir $v_i = v|_{\Omega_i}$. Dann gilt $v_i \in H^1(\Omega_i)$. Sei

$$(2.15) \quad V := \{ v \in H^1(\Omega) \mid \sqrt{k_1 \lambda_2} v_1|_S = \sqrt{k_2 \lambda_1} v_2|_S \}$$

(vergl. LIONS [30]) und $H := L^2(\Omega)$. Hierbei ist die Restriktion auf S im Sinne von Satz 1.1 zu verstehen. Wir versehen V mit dem Skalarprodukt von $H^1(\Omega_i)$. Dann ist V ein Hilbert-Raum, da V ein abgeschlossener Teilraum von $H^1(\Omega)$ ist. Oberdies liegt V dicht in $L^2(\Omega)$, da $\mathcal{D}(\Omega_i) \subset V$ dicht in $L^2(\Omega_i)$ liegt. Sei

$$(2.16) \quad k(x) := \begin{cases} k_1, & \text{falls } x \in \Omega_1 , \\ k_2, & \text{falls } x \in \Omega_2 , \end{cases}$$

und

$$(2.17) \quad a(u, v) := \sum_{i=1}^n (k \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i}) .$$

Da diese Bilinearform stetig auf $V \times V$ ist und nicht von t abhängt, folgt sofort

(1.32) und (1.33). Eine einfache Rechnung zeigt die Gültigkeit von (1.34). Man beachte dabei, daß die Konstanten k_1, k_2 aus physikalischen Gründen positiv sind. Wir wählen $U := L^2(Q_2)$ und setzen

$$(2.18) \quad (Bu)(t)(x) := \begin{cases} u(t, x), & \text{falls } (t, x) \in Q_2, \\ 0, & \text{falls } (t, x) \in Q_1. \end{cases}$$

Dann gilt $Bu \in L^2(0, T; L^2(\Omega)) \subset L^2(0, T; V^*)$ und B ist stetig. Weiter setzen wir $e := 0$. Wir müssen zeigen, daß in dieser Situation die Interpretation von (2.6) gerade das System (2.8) - (2.10), (2.11'), (2.12') liefert. Für alle $t \in]0, T[$ und $i = 1, 2$ sei

$$(2.19) \quad \frac{\partial y_i}{\partial t}(t) = \frac{dy}{dt}(t)|_{H_0^{-1}(\Omega_i)}.$$

Dabei fassen wir $H_0^{-1}(\Omega_i)$ als Teilraum von V auf, indem wir $v \in H_0^{-1}(\Omega_i)$ durch Null auf Ω fortsetzen. $\frac{\partial y_i}{\partial t}(t)$ liegt in $H^{-1}(\Omega_i)$ und $\frac{\partial y_i}{\partial t}$ kann daher in naheliegender Weise als Distribution auf Q_i angesehen werden. Aus (2.6) folgt sofort für fast alle $t \in]0, T[$

$$(2.20) \quad \frac{\partial y_1}{\partial t}(t) - k_1 \Delta y_1(t) = 0,$$

$$(2.21) \quad \frac{\partial y_2}{\partial t}(t) - k_2 \Delta y_2(t) = u(t).$$

Also ist (2.6) hier mit (2.20), (2.21) und

$$(2.22) \quad (v, \frac{dy}{dt}(t)) + a(v, y(t)) = \sum_{i=1}^2 (v_i, \frac{\partial y_i}{\partial t}(t) - k_i \Delta y_i(t)) \quad \forall t \in]0, T[, \forall v \in V$$

äquivalent.

Wir nehmen an, daß rein formal die folgende Green'sche Formel gilt:

$$(2.23) \quad \begin{aligned} (v, \frac{dy}{dt}(t)) + a(v, y(t)) &= (v_1, \frac{\partial y_1}{\partial t}(t) - k_1 \Delta y_1(t)) \\ &+ k_1 \int_{\partial\Omega_1} v_1 \frac{\partial y_1}{\partial n}(t) d\Gamma + (v_2, \frac{\partial y_2}{\partial t}(t) - k_2 \Delta y_2(t)) \\ &+ k_2 \int_{\partial\Omega_2} v_2 \frac{\partial y_2}{\partial n}(t) d\Gamma. \end{aligned}$$

Dann ist, ebenfalls rein formal, (2.22) äquivalent mit

$$(2.24) \sum_{i=1}^2 k_i \int_{\Gamma_i} v_i \frac{\partial y_i}{\partial n}(t) d\Gamma + \int_S v_1 (k_1 \frac{\partial y_1}{\partial n}(t) + k_2 \frac{\gamma_{k_1, \lambda_2}}{\gamma_{k_2, \lambda_1}} \frac{\partial y_2}{\partial n}(t)) d\Gamma = 0 \quad \forall v \in V.$$

Wählt man erst Funktionen v mit $v_1=0$, dann mit $v_2=0$ und schließlich mit $v_1|_{\Gamma_1} = v_2|_{\Gamma_2} = 0$, so sieht man, daß (2.24) mit (2.10), (2.12') äquivalent ist.

Es fragt sich natürlich, ob die Green'sche Formel (2.23) gilt. Dabei hat man zunächst die Normalableitungen geeignet zu definieren. Vermutlich ist dies möglich (vergl. ähnlich gelagerte Fälle in LIONS [25], S. 24, S. 122), erfordert aber die recht komplizierte Definition der Sobolev-Räume $H^s(\Gamma)$, $s \in \mathbb{R}$. Die Normalableitungen sind dann gewisse stetige Linearformen auf $H^s(\Gamma)$ (mit geeignetem $s \geq 0$). Wir wollen diesen Fragen nicht weiter nachgehen, sondern begnügen uns damit, die Gleichung (2.22) selbst als geeignete Verallgemeinerung der Rand- und Übergangsbedingungen (2.10) und (2.12') anzusehen.

Man kann analog zu Satz 1.1 für Elemente aus $L^2(0, T; H^1(\Omega_1))$ die Restriktion auf $]0, T[\times \partial\Omega_1$ definieren. Die Restriktionsabbildung ist stetig und hat ihre Werte in $L^2(]0, T[\times \partial\Omega_1)$ (vergl. LIONS-MAGENES [32], Kap. IV). Wir können daher

$$(2.25) \quad \mathcal{B} := L^2(]0, T[\times S), \quad Cy := y_1|_{]0, T[\times S}, \quad f := \| \cdot \|_{\mathcal{B}}^2$$

definieren. Weiter sei

$$(2.26) \quad U_{ad} := \{u \in L^2(Q_2) \mid |u(t, x)| \leq m \quad \forall (t, x) \in Q_2\}.$$

U_{ad} ist abgeschlossen und beschränkt in $L^2(Q_2)$.

(b) Wir ändern dieses Beispiel so ab, daß f nicht auf einem Hilbert-Raum definiert ist. Dazu nehmen wir an, daß die Temperatur in ganz Ω_1 im Sinne der L_1 -Norm kontrolliert werden soll. Sei

$$(2.27) \quad \mathcal{B} := L^1(]0, T[\times \Omega_1), \quad Cy := y_1, \quad f := \| \cdot \|_{\mathcal{B}}.$$

§3. Konvexe Optimierung und Dualität

Zu jedem konvexen Optimierungsproblem kann man ein sogenanntes "duales" Optimierungsproblem angeben. Mit diesem Dualproblem lassen sich in sehr allgemeiner und eleganter Weise Charakterisierungs- und Existenzfragen für Optimierungsprobleme behandeln. Überdies erweist es sich auch bei Konvergenzuntersuchungen diskretisierter Optimierungsprobleme als hilfreich, wie wir im nächsten Kapitel sehen werden. Aus diesen Gründen sollen in aller Kürze die wichtigsten Ergebnisse aus der Dualitätstheorie dargestellt werden.

Wir beginnen mit einigen Definitionen. Sei $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ und sei X ein topologischer linearer Raum. Eine Funktion $f: X \rightarrow \bar{\mathbb{R}}$ heißt konvex, wenn ihr Epigraph

$$(3.1) \quad \text{epif} := \{(x, \mu) \in X \otimes \mathbb{R} \mid \mu \geq f(x)\}$$

konvex ist. Der effektive Definitionsbereich von f wird durch

$$(3.2) \quad \text{dom} f := \{x \in X \mid f(x) < +\infty\}$$

definiert. f heißt eigentlich, wenn f nicht identisch $+\infty$ und nirgends $-\infty$ ist. Seien X und Y lokalkonvexe topologische lineare Räume, die vermöge einer Bilinearform $\langle \cdot, \cdot \rangle$ ein topologisches Dualsystem bilden. Dann wird das zu f konjugierte Funktional $f^*: Y \rightarrow \bar{\mathbb{R}}$ durch

$$(3.3) \quad f^*(y) := \sup_{x \in X} [\langle x, y \rangle - f(x)]$$

definiert. f^* ist ebenfalls konvex. Das Subdifferential $\partial f(x_0)$ von f an einer Stelle $x_0 \in X$ ist durch

$$(3.4) \quad \partial f(x_0) := \{y \in Y \mid f(x) \geq f(x_0) + \langle x - x_0, y \rangle \quad \forall x \in X\}$$

gegeben. Die Elemente von $\partial f(x_0)$ heißen Subgradienten. Entsprechende Definitionen führt man auch für konkave Funktionen ein (vergl. ROCKAFELLAR[35]). Sei $C \subset X$ konvex. Die Indikatorfunktion bezüglich C wird durch

$$(3.5) \quad \delta(x \mid C) := \begin{cases} 0, & \text{falls } x \in C, \\ +\infty & \text{falls } x \notin C. \end{cases}$$

erklärt. Schließlich bezeichne $\text{int } C$ bzw. $\text{ri } C$ das Innere bzw. das relativ

Innere von C (d.h. das Innere bezüglich der affinen Hülle von C).

Seien $X_1, Y_1; X_2, Y_2$ zwei Paare lokalkonvexer topologischer linearer Räume, die vermöge zweier Bilinearformen $\langle \cdot, \cdot \rangle_1, \langle \cdot, \cdot \rangle_2$ jeweils ein topologisches Dualsystem bilden. Weiter seien eine konvexe Funktion $f: X_1 \rightarrow \bar{\mathbb{R}}$ und zu jedem $u \in X_2$ eine konvexe Funktion $F_u: X_1 \rightarrow \bar{\mathbb{R}}$ derart gegeben, daß $F_0 = f$ gilt. Ist dann die Abbildung $(x, u) \mapsto (F_u)(x)$ konvex, so heißt F eine (zu f gehörige) konvexe Bifunktion. Als Optimierungsproblem betrachten wir

(P) Minimiere $(F_0)(x)$ über alle $x \in X_1$.

Nebenbedingungen sind implizit dadurch gegeben, daß F_0 den Wert $+\infty$ annehmen kann. Die zu F adjungierte Bifunktion wird für $x^* \in Y_1, u^* \in Y_2$ durch

$$(3.6) \quad (F^*x^*)(u^*) := \inf_{x \in X_1, u \in X_2} [(F_u)(x) - \langle x, x^* \rangle_1 + \langle u, u^* \rangle_2]$$

definiert. F^* ist konkav. Als Dualproblem setzt man jetzt

(D) Maximiere $(F^*0)(u^*)$ über alle $u^* \in Y_2$.

Weiter benötigen wir die zu F und F^* gehörenden "Störfunktionen"

$$(3.7) \quad \varphi_F(u) := \inf_{x \in X_1} (F_u)(x), \quad \varphi_{F^*}(x^*) := \sup_{u^* \in Y_2} (F^*x^*)(u^*).$$

φ_F ist konvex und φ_{F^*} konkav.

Lemma 3.1. $u_0^* \in \partial(-\varphi_F)(0)$ gilt genau dann, wenn

$$(3.8) \quad (F^*0)(u_0^*) = \max_{u^* \in Y_2} (F^*0)(u^*) = \inf_{x \in X_1} (F_0)(x).$$

Der Beweis dieses Lemmas besteht aus einer einfachen Rechnung, bei der man die für alle $x \in X_1, u^* \in Y_2$ gültige Ungleichung

$$(3.9) \quad (F^*0)(u^*) \leq (F_0)(x)$$

verwendet. Der wesentliche Schritt bei der Aufstellung der Gleichung (3.8) besteht also darin, die Subdifferenzierbarkeit von $-\varphi_F$ im Nullpunkt nachzuweisen (d.h. die Gültigkeit von $\partial(-\varphi_F)(0) \neq \emptyset$ zu sichern). Man kann dazu

zeigen, daß eine konvexe Funktion f an einer Stelle $x_0 \in \text{dom } f$ subdifferenzierbar ist, wenn $\text{epi } f$ und der Punkt $(x_0, f(x_0))$ durch eine nichtsenkrechte abgeschlossene Hyperebene trennbar sind. Dann folgt leicht mit einem bekannten Trennungssatz der

Satz 3.1. Gilt $0 \in \text{ri}(\text{dom } \varphi_F)$ und $\text{ri}(\text{epi } \varphi_F) \neq \emptyset$, so folgt

$$(3.10) \quad \inf_{x \in X_1} (F0)(x) = \max_{u^* \in Y_2} (F^* 0)(u^*).$$

Bemerkung: Falls $F^* 0 \neq -\infty$, folgt aus diesem Satz die Existenz einer Optimallösung im dualen Problem.

Es liegt nahe, dem Problem (D) wieder ein duales Problem zuzuordnen, indem man die zu F^* adjungierte Bifunktion F^{**} bildet. Ist F eigentlich und nach unten halbstetig, so ergibt sich wieder (P). Dies folgt daraus, daß für eine eigentliche nach unten halbstetige Funktion f die Gleichung $f^{**} = f$ gilt. Führt man obige Überlegungen mit F^* anstelle von F durch (und beachtet, daß F^* konkav ist), so erhält man den folgenden Satz.

Satz 3.1* Ist F eine eigentliche, nach unten halbstetige konvexe Bifunktion, so folgt aus $0 \in \text{ri}(\text{dom } \varphi_{F^*})$ und $\text{ri}(\text{epi } \varphi_{F^*}) \neq \emptyset$ die Gleichung

$$(3.11) \quad \min_{x \in X_1} (F0)(x) = \sup_{u^* \in Y_2} (F^* 0)(u^*).$$

Bemerkung: Falls $F0 \neq +\infty$, so folgt aus diesem Satz die Existenz einer Optimallösung im primalen Problem.

Weiter erhält man durch das Dualproblem Charakterisierungssätze.

Satz 3.2. Unter den Voraussetzungen von Satz 3.1 ist x für (P) genau dann optimal, wenn es ein $u^* \in Y_2$ gibt mit

$$(3.12) \quad (F0)(x) = (F^* 0)(u^*).$$

Beweis: Die Notwendigkeit von (3.12) folgt sofort aus Satz 3.1. Andererseits ist (3.12) auch ohne die Voraussetzung von Satz 3.2 hinreichend, wie man (3.9) entnimmt. $\square$

Satz 3.2*. Unter den Voraussetzungen von Satz 3.1* ist u^* für (D) genau dann optimal, wenn es ein $x \in X_1$ mit (3.12) gibt.

§4. Dualität bei parabolischen Kontrollproblemen

4.1. Die Herleitung des Dualproblems

Mit den Ergebnissen aus §3 sind wir jetzt in der Lage, zu dem Kontrollproblem (P) aus §2.1 ein Dualproblem aufzustellen. Sei

$$(4.1) \quad \begin{aligned} X_1 &:= L^2(0, T; V) \oplus U, & Y_1 &:= L^2(0, T; V^*) \oplus U, \\ X_2 &:= L^2(0, T; V^*) \oplus B, & Y_2 &:= L^2(0, T; V) \oplus B^*. \end{aligned}$$

Wir identifizieren U^* mit U , versehen B^* mit der schwach*-Topologie und definieren für $(y, u) \in X_1, (y^*, u^*) \in Y_1, (w, v) \in X_2$ und $(w^*, v^*) \in Y_2$:

$$\begin{aligned} \langle (y, u), (y^*, u^*) \rangle_1 &:= \langle y, y^* \rangle + \langle u, u^* \rangle_U, \\ \langle (w, v), (w^*, v^*) \rangle_2 &:= \langle w^*, w \rangle + \langle v, v^* \rangle_{B^*}. \end{aligned}$$

Hierbei sei $\langle v, v^* \rangle_{B^*} := v^*(v)$ und

$$(4.2) \quad [z, z^*] := \int_0^T (z(t), z^*(t)) dt$$

für alle $z \in L^2(0, T; V)$, $z^* \in L^2(0, T; V^*)$. Für jedes $w \in L^2(0, T; V^*)$ definieren wir die Menge

$$(4.3) \quad K_w := \left\{ (y, u) \in X_1 \mid \frac{dy}{dt} + A(t)y = -w + e + Bu, y(0) = y_0, u \in U_{ad} \right\}$$

und führen die Bifunktion

$$(4.4) \quad [F(w, v)](y, u) := f(Cy + v - z_d) + \langle u, Nu \rangle_U + \delta(y, u \mid K_w)$$

ein. Schließlich bezeichne für alle $t \in]0, T[$ $A^*(t)$ den zu $A(t)$ adjungierten Operator.

Satz 4.1. Das zur Bifunktion F aus (4.4) gehörende Dualproblem lautet wie folgt.

(D) Maximiere

$$(4.5) \quad -f^*(v^*) + \min_{u \in U_{ad}} \langle u, Nu + B^* w^* \rangle_U - \langle z_d, v^* \rangle_{B^*} + [w^*, e] + (y_0, w^*(0))$$

über alle $w^* \in L^2(0, T; V)$ und $v^* \in B^*$ unter den Nebenbedingungen

$$(4.6) \quad -\frac{dw^*}{dt} + A^*(t)w^* = C^* v^*,$$

$$(4.7) \quad w^*(T) = 0.$$

Beweis: Wir müssen gemäß Gleichung (3.6) die zu F adjungierte Bifunktion berechnen.

$$\begin{aligned}
 [F(y^*, u^*)](w^*, v^*) &= \inf \{ f(Cy + v - z_d) + \langle u, Nu \rangle_U - [y, y^*] - \langle u, u^* \rangle_U + [w^*, w] \\
 &\quad + \langle v, v^* \rangle_B \mid (y, u) \in X_1, (w, v) \in X_2, \frac{dy}{dt} + A(t)y = -w + e + Bu, \\
 &\quad y(0) = y_0, u \in U_{ad} \} \\
 &= \inf \{ f(Cy + v - z_d) + \langle u, Nu \rangle_U - [y, y^*] - \langle u, u^* \rangle_U + \langle v, v^* \rangle_B \\
 &\quad + [w^*, e + Bu - \frac{dy}{dt} - A(t)y] \mid y \in W(0, T), v \in B, y(0) = y_0, u \in U_{ad} \} \\
 &= \inf \{ f(v) + \langle u, Nu \rangle_U - [y, y^*] - \langle u, u^* \rangle_U + \langle v - Cy + z_d, v^* \rangle_B \\
 &\quad + [w^*, e + Bu - \frac{dy}{dt} - A(t)y] \mid y \in W(0, T), v \in B, y(0) = y_0, u \in U_{ad} \} \\
 &= \inf \{ f(v) + \langle v, v^* \rangle_B + \langle u, -u^* + Nu + B^* w^* \rangle_U - [w^*, \frac{dy}{dt} + A(t)y] \\
 &\quad - [y, C^* v^* + y^*] + \langle z_d, v^* \rangle_B + [w^*, e] \mid y \in W(0, T), v \in B, \\
 &\quad y(0) = y_0, u \in U_{ad} \}.
 \end{aligned}$$

Wir setzen

$$M(y^*, w^*, v^*) := \inf \{ -[w^*, \frac{dy}{dt} + A(t)y] - [y, C^* v^* + y^*] \mid y \in W(0, T), y(0) = y_0 \}.$$

Dann gilt

$$\begin{aligned}
 [F^*(y^*, u^*)](w^*, v^*) &= -f^*(-v^*) + \inf_{u \in U_{ad}} \langle u, -u^* + Nu + B^* w^* \rangle_U + M(y^*, w^*, v^*) \\
 &\quad + \langle z_d, v^* \rangle_B + [w^*, e].
 \end{aligned}$$

Mit $z^* := -A^*(t)w^* - C^*v^* - y^*$ ergibt sich

$$M(y^*, w^*, v^*) = \inf \{ -[w^*, \frac{dy}{dt}] + [y, z^*] \mid y \in W(0, T), y(0) = y_0 \}.$$

Wähle ein $y_1 \in W(0, T)$ mit $y_1(0) = y_0$. Dann folgt

$$M(y^*, w^*, v^*) = \inf \{ -[w^*, \frac{dy}{dt}] + [y, z^*] \mid y \in W(0, T), y(0) = 0 \} - [w^*, \frac{dy_1}{dt}] + [y_1, z^*].$$

$M(y^*, w^*, v^*)$ ist genau dann endlich, wenn

$$(4.8) \quad [w^*, \frac{dy}{dt}] = [y, z^*] \quad \text{für alle } y \in W(0, T) \text{ mit } y(0) = 0.$$

Wir setzen

$$D(\wedge) := \{ y \in L^2(0, T; V) \mid \frac{dy}{dt} \in L^2(0, T; V), y(0) = 0 \}$$

und definieren einen unbeschränkten Operator $\wedge : D(\wedge) \rightarrow L^2(0, T; V)$ durch

$\wedge := \frac{d}{dt}$. Der im Sinne der unbeschränkten Operatoren zu $\wedge$ adjungierte Operator $\wedge^* : D(\wedge^*) \rightarrow L^2(0, T; V^*)$ ist durch $\wedge^* = -\frac{d}{dt}$ gegeben, wobei

$$D(\wedge^*) = \left\{ y \in L^2(0, T; V^*) \mid \frac{dy}{dt} \in L^2(0, T; V^*), y(T) = 0 \right\},$$

vergleiche hierzu LIONS - MAGENES [31], Kap. III, §§ 1, 4.

Gleichung (4.8) gilt insbesondere für alle $y \in D(\wedge)$. Daher ist das lineare Funktional $y \mapsto [w^*, \frac{dy}{dt}]$ stetig auf $D(\wedge)$. Dann folgt $w^* \in D(\wedge^*)$ und

$$-[y, \frac{dw^*}{dt}] = [y, z^*] \quad \forall y \in D(\wedge).$$

Da $D(\wedge)$ dicht in $L^2(0, T; V)$ liegt, haben wir also

$$(4.9) \quad z^* = -\frac{dw^*}{dt}, \quad w^* \in D(\wedge^*).$$

$D(\wedge)$ liegt dicht in der Menge aller $y \in W(0, T)$ mit $y(0) = 0$, wenn man diesen linearen Teilraum von $W(0, T)$ mit der von $W(0, T)$ induzierten Topologie versieht. Gilt (4.8) für alle $y \in D(\wedge)$, so also auch für alle $y \in W(0, T)$ mit $y(0) = 0$. Damit ist gezeigt, daß (4.8) mit (4.9) äquivalent ist.

Wegen $y_1 \in W(0, T)$ und $w^* \in D(\wedge^*) \cap L^2(0, T; V)$ können wir partiell integrieren (vergl. LIONS - MAGENES [31], Kap. III, § 4.6), also

$$(4.10) \quad -[w^*, \frac{dy_1}{dt}] - [y_1, \frac{dw^*}{dt}] = (y_1(0), w^*(0)).$$

Folglich ergibt sich als adjungierte Bifunktion

$$(4.11) [F^*(y^*, u^*)](w^*, v^*) = \begin{cases} -f^*(-v^*) + \inf_{u \in U_{ad}} \langle u, -u^* + Nu + B^* w^* \rangle_U + (y_0, w^*(0)) + \langle z_d, v^* \rangle_{\beta} \\ \quad + [w^*, e], \\ \text{falls } \frac{dw^*}{dt} = A^*(t)w^* + C^* v^* + y^* \text{ und } w^*(T) = 0, \\ -\infty, \text{ sonst.} \end{cases}$$

Durch einen einfachen Kompaktheitsschluß läßt sich leicht zeigen, daß unter der Voraussetzung (2.2) in (4.11) "inf" durch "min" ersetzt werden kann. Damit ist alles bewiesen. $\square$

Wir weisen darauf hin, daß mit dem Dualproblem aus Satz 4.1 zumindest prinzipiell untere Schranken für das Infimum des primalen Kontrollproblems gefunden werden können. Nach Ungleichung (3.9) liefert der Wert der Zielfunktion (4.5) für ein beliebiges zulässiges Paar (w^*, v^*) eine solche untere Schranke.

4.2. Der Dualitätssatz

Wir bezeichnen das Infimum (Minimum) im primalen Problem mit $\inf (P)$ ($\min(P)$) und verwenden eine entsprechende Abkürzung für das Dualproblem.

Satz 4.2. Unter den in § 2.1 gemachten Voraussetzungen existiert sowohl für das Problem (P) aus § 2.1 als auch für das Problem (D) aus § 4.1 eine Optimallösung und es gilt

$$(4.12) \quad \min (P) = \max (D).$$

Beweis: Wir haben die Voraussetzungen der Sätze 3.1, 3.1* nachzuweisen.

(a) Zu jeder rechten Seite von (2.4) erhält man für festes y_0 genau eine Lösung y der Differentialgleichung. Der so konstruierte Operator

$$(4.13) \quad P : L^2(0,T;V^*) \rightarrow W(0,T)$$

ist stetig (vergl. LIONS-MAGENES [31], Kap III). Wähle ein festes $u \in U_{ad}$. Dann gilt für alle $(w,v) \in X_2$

$$(4.14) \quad \begin{aligned} \varphi_F(w,v) &\leq [F(w,v)](P(-w+e+Bu), u) \\ &= f(C(P(-w+e+Bu)+v-z_d)) + \langle u, Nu \rangle_U. \end{aligned}$$

Also folgt

$$\text{dom } \varphi_F = L^2(0,T;V^*) \oplus \mathcal{B}.$$

Da die rechte Seite von (4.14) stetig in (w,v) auf X_2 ist, wird φ_F durch eine stetige Funktion majorisiert. Hieraus ergibt sich unmittelbar $\text{int}(\text{epi } \varphi_F) \neq \emptyset$.

(b) Analog wie in (a) gibt es zu jeder rechten Seite von (4.6) genau eine Lösung w^* mit $w^*(T) = 0$ und man erhält einen stetigen Operator

$$(4.15) \quad \bar{P} : L^2(0,T;V^*) \rightarrow W(0,T).$$

Es gilt für eine konvexe Funktion $f^* \equiv +\infty$ genau dann, wenn $\liminf f \equiv -\infty$ (vergl. MOREAU [33]). Folglich gibt es ein $v_0^* \in \mathcal{B}^*$ mit $f^*(v_0^*) < +\infty$. Wir definieren

$$(4.16) \quad m(y^*, u^*) := \min_{u \in U_{ad}} \langle u, -u^* + Nu - B^*(\bar{P}(C^* v_0^* + y^*)) \rangle_U$$

und erhalten hiermit für alle $(y^*, u^*) \in Y_1$ die Ungleichung

$$(4.17) \quad \varphi_{F^*}(y^*, u^*) \geq -f^*(-v_0^*) + m(y^*, u^*) - (y_0^*, \bar{P}(C^* v_0^* + y^*)) (0) + \langle z_d, v_0^* \rangle_{\mathcal{B}} - [\bar{P}(C^* v_0^* + y^*), e].$$

1. Fall: U_{ad} ist beschränkt. Dann gibt es eine Kugel K mit Radius r um den Nullpunkt von U , so daß $U_{ad} \subset K$. Da N positiv semidefinit ist, gilt für alle $(y^*, u^*) \in Y_1$

$$(4.18) \quad \begin{aligned} m(y^*, u^*) &\geq \min_{u \in U_{ad}} \langle u, -u^* - B^*(\bar{P}(C^* v_0^* + y^*)) \rangle_U \\ &\geq -r \|u^* + B^*(\bar{P}(C^* v_0^* + y^*))\|_U. \end{aligned}$$

2. Fall: N ist koerzitiv. Dann gilt für beliebige $u^*, u_1^* \in U$

$$\begin{aligned} \langle u, Nu + u_1^* \rangle_U &\geq \nu \|u\|_U^2 + \langle u, u_1^* \rangle_U \\ &= \nu \|u\|_U^2 + \frac{1}{2\nu} \|u_1^*\|_U^2 - \frac{1}{4\nu} \|u_1^*\|_U^2, \end{aligned}$$

also für alle $(y^*, u^*) \in Y_1$

$$(4.19) \quad m(y^*, u^*) \geq -\frac{1}{4\nu} \|u^* + B^*(\bar{P}(C^* v_0^* + y^*))\|_U^2.$$

In beiden Fällen gilt also

$$\text{dom } \varphi_{F^*} = L^2(0, T; V^*) \oplus U$$

und φ_{F^*} ist durch eine stetige Funktion nach unten beschränkt. Daher folgt $\text{int}(\text{epi } \varphi_{F^*}) \neq \emptyset$.

Berücksichtigt man, daß U_{ad} abgeschlossen ist, so sieht man leicht, daß die Niveaumengen $\{(w, v, y, u) \in X_2 \oplus X_1 \mid [F(w, v)](y, u) \leq \alpha\}$, $\alpha \in \mathbb{R}$, von F abgeschlossen sind. Daher ist F nach unten halbstetig. $\square$

4.3. Beispiele

1. Quadratische Probleme. Sei $\mathcal{B} = \mathcal{H}$ ein Hilbert-Raum und $f(v) := \|v\|_{\mathcal{H}}^2$ für alle $v \in \mathcal{H}$. Mit einer analogen Überlegung, wie sie zur Herleitung von (4.19) nötig war, überlegt man sich, daß

$$(4.20) \quad f^*(v^*) = \frac{1}{4} \|v^*\|_{\mathcal{H}}^2.$$

2. Minimumnormprobleme. Sei $f(v) := \|v\|_{\mathcal{B}}$. Es ist klar, daß $f^*(v^*)$ genau dann endlich ist, wenn

$$(4.21) \quad \langle v, v^* \rangle_{\mathcal{B}} - \|v\|_{\mathcal{B}} \leq 0 \quad \forall v \in \mathcal{B}.$$

(4.21) ist äquivalent mit

$$(4.22) \quad \|v^*\|_{\mathcal{B}^*} \leq 1.$$

In diesem Fall verschwindet also $f^*(v^*)$ in (4.5), dafür tritt zu den Nebenbedingungen die Ungleichung (4.22) hinzu.

3. Das Transmissions-Kontrollproblem. Wir wollen zu dem Wärmeleitungsproblem aus § 2.2 das Dualproblem explizit berechnen. Es gilt hier

$$(4.23) \quad \begin{aligned} s := \min_{u \in U_{ad}} \langle u, Nu + B^* w^* \rangle_U &= \min_{u \in U_{ad}} [w^*, Bu] \\ &= \min \left\{ \int_0^T \int_{\Omega_2} w_2^* u \, dx \, dt \mid u \in L^\infty(Q_2), |u| \leq m \right\}. \end{aligned}$$

Bezeichne $\|\cdot\|_\infty$ die Norm von $L^\infty(Q_2)$ und $\|\cdot\|_1$ die Norm von $L^1(Q_2)$. Mit der Ungleichung

$$\|u\|_\infty \|w_2^*\|_1 \leq \int_0^T \int_{\Omega_2} u w_2^* \, dx \, dt$$

folgt aus (4.23) sofort

$$(4.24) \quad s = -m \int_0^T \int_{\Omega_2} |w_2^*| \, dx \, dt.$$

Die Variationsform der adjungierten Differentialgleichung (4.6) ist gegeben durch

$$(4.25) \quad -(v, \frac{dw^*}{dt}(t)) + a(t, w^*(t), v) = (v, (C^* v^*)(t)) \quad \forall t \in]0, T[, \quad \forall v \in V.$$

Man kann (4.25) analog wie in § 2.2 interpretieren und erhält dann als Dualproblem: Maximiere

$$(4.26) \quad - \int_0^T \int_{\Omega} v^* \left(\frac{1}{4} v^* + z_d \right) d\Gamma \, dt - m \int_0^T \int_{\Omega_2} |w_2^*| \, dx \, dt + \sum_{i=1}^2 \int_{\Omega_i} y_{0i} w_i^*(0) \, dx$$

über alle $w^* \in L^2(0, T; V)$, $v^* \in L^2(]0, T[\times S)$ unter den Nebenbedingungen ($i=1, 2$)

$$(4.27) \quad \frac{\partial w_i^*}{\partial t}(t) + k_i \Delta w_i^*(t) = 0,$$

$$(4.28) \quad w_i^*(T) = 0,$$

$$(4.29) \quad -(v, \frac{dw^*}{dt}(t)) + a(t, w^*(t), v) = \sum_{i=1}^2 (v_i, -\frac{\partial w_i^*}{\partial t} - k_i \Delta w_i^*(t)) + \int_S v_{|S}(t) v^*(t) d\Gamma$$

$$\forall t \in]0, T[, \forall v \in V.$$

Man kann (4.29) wieder mit Hilfe einer Green'schen Identität umformen. In diesem Sinne ist (4.29) äquivalent mit

$$(4.30) \quad \frac{\partial w_i^*}{\partial n} = 0 \quad \text{auf } \Gamma_i,$$

$$(4.31) \quad \lambda_1 \sqrt{k_1 \lambda_2} \frac{\partial w_1^*}{\partial n} + \lambda_2 \sqrt{k_2 \lambda_1} \frac{\partial w_2^*}{\partial n} = \frac{\lambda_1 \sqrt{\lambda_2}}{\sqrt{k_1}} v^* \quad \text{auf } S.$$

§5. Charakterisierungen optimaler Steuerungen

5.1. Das allgemeine Kontrollproblem

Wir wenden Satz 3.2 an, um Optimalitätskriterien zu gewinnen. Demnach sind also wegen Satz 4.2 für das Primalproblem (P) zulässige Funktionen y, u genau dann optimal, wenn es ein $w^* \in L^2(0, T; V^*)$ und ein $v^* \in \mathcal{B}^*$ mit (4.6), (4.7) gibt, so daß

$$(5.1) \quad f(Cy - z_d) + \langle u, Nu \rangle_U \leq -f^*(v^*) + \langle w, Nw + B^* w^* \rangle_U - \langle z_d, v^* \rangle_{\mathcal{B}} \\ + [w^*, e] + (y_0, w^*(0)) \quad \forall w \in U_{ad}.$$

Um dies einzusehen, beachte man (3.9). Wegen (2.4) ist (5.1) äquivalent mit

$$f(Cy - z_d) + \langle u, Nu + B^* w^* \rangle_U \leq -f^*(v^*) + \langle w, Nw + B^* w^* \rangle_U - \langle z_d, v^* \rangle_{\mathcal{B}} \\ + [w^*, \frac{dy}{dt} + A(t)y] + (y_0, w^*(0)) \quad \forall w \in U_{ad}$$

$$(5.2) \quad = -f^*(v^*) + \langle w, Nw + B^* w^* \rangle_U + \langle Cy - z_d, v^* \rangle_{\mathcal{B}} \\ \forall w \in U_{ad}.$$

Um (5.2) zu erhalten, haben wir partiell integriert, was genau wie in §4.1 möglich war, und dann Gleichung (4.6) angewendet. Setzen wir in (5.2) speziell $w = u$, so folgt

$$(5.3) \quad \sup_{v \in \mathcal{B}} [\langle v, v^* \rangle_{\mathcal{B}} - f(v)] = \langle Cy - z_d, v^* \rangle_{\mathcal{B}} - f(Cy - z_d).$$

Unter Verwendung von (5.3) ergibt sich aus (5.2)

$$(5.4) \quad \langle u, Nu + B^* w^* \rangle_U \leq \langle w, Nw + B^* w^* \rangle_U \quad \forall w \in U_{ad}.$$

Da U_{ad} konvex ist, ist diese Bedingung äquivalent mit

$$(5.5) \quad \langle u, Nu + B^* w^* \rangle_U \leq \langle u + \lambda(w - u), N(u + \lambda(w - u)) + B^* w^* \rangle_U \quad \forall w \in U_{ad}, \forall \lambda \in [0, 1]$$

(5.5) gilt genau dann, wenn

$$(5.6) \quad -\lambda^2 \langle w - u, N(w - u) \rangle_U \leq \lambda \langle w - u, 2Nu + B^* w^* \rangle_U \quad \forall w \in U_{ad}, \quad \forall \lambda \in [0, 1].$$

Da N positiv semidefinit ist, bedeutet (5.6) also nichts anderes als

$$(5.7) \quad 0 \leq \langle w - u, 2Nu + B^* w^* \rangle_U \quad \forall w \in U_{ad}.$$

Damit haben wir den folgenden Satz bewiesen.

Satz 5.1. Ein Funktionenpaar $(y, u) \in L^2(0, T; V) \oplus U$ ist genau dann Optimallösung für das Kontrollproblem (P) aus § 2.1, wenn es zulässig ist und ein für das Dualproblem (D) aus § 4.1 zulässiges Funktionenpaar $(w^*, v^*) \in L^2(0, T; V^*) \oplus \mathbb{R}^1$ existiert, so daß

$$(5.8) \quad v^* \in \partial f(Cy - z_d),$$

$$(5.9) \quad \langle u, 2Nu + B^* w^* \rangle_U \leq \langle w, 2Nu + B^* w^* \rangle_U \quad \forall w \in U_{ad}.$$

Bemerkung: (5.8) und (5.9) bezeichnen wir als Optimalitätsbedingungen. Man beachte aber, daß zur Bestimmung von y, u, v^*, w^* zusätzlich die Bedingungen $u \in U_{ad}$, (2.4), (2.5), (4.6), (4.7) heranzuziehen sind.

5.2. Beispiele

1. Quadratische Probleme. Seien f und $\mathcal{B}$ wie im ersten Beispiel von § 4.3 gegeben. Man errechnet leicht, daß (5.8) dann mit

$$(5.10) \quad v^* = \frac{1}{2} (Cy - z_d)$$

äquivalent ist. Man kann v^* in (4.6) einsetzen und auf diese Weise eliminieren.

2. Minimumnormprobleme. Seien f und $\mathcal{B}$ wie im zweiten Beispiel von § 4.3 gegeben. (5.8) ist mit (5.3) äquivalent, also gilt hier (5.8) genau dann, wenn

$$(5.11) \quad \langle Cy - z_d, v^* \rangle_{\mathcal{B}} = \|Cy - z_d\|_{\mathcal{B}}$$

und $\|v^*\|_{\mathcal{B}^*} \leq 1$. Setzen wir $Cy - z_d \neq 0$ voraus, so ist daher (5.8) mit (5.11) und

$$(5.12) \quad \|v^*\|_{\mathcal{B}^*} = 1$$

äquivalent.

3. Das Transmissions-Kontrollproblem. Wir betrachten erneut das Wärmeleitungsproblem aus § 2.2. Wegen (4.24) ist (5.9) hier äquivalent mit

$$(5.13) \quad \int_0^T \int_{\Omega_2} [u w_2^* + m |w_2^*|] dx dt = 0.$$

Aus $u \in U_{ad}$ folgt $-u w_2^* \leq |u| |w_2^*| \leq m |w_2^*|$. Also gilt (5.13) genau dann, wenn

$$(5.14) \quad u w_2^* + m |w_2^*| = 0.$$

Kann man garantieren, daß w_2^* höchstens auf einer Menge vom Maße Null verschwindet, so bedeutet (5.14) nichts anderes als

$$(5.15) \quad |u| = m.$$

Es gilt dann für u das Bang-Bang-Prinzip. Wir wollen der Frage, ob w_2^* die genannte Eigenschaft hat, nicht weiter nachgehen.

Schließlich analysieren wir noch Bedingung (5.8) für das modifizierte Wärmeleitungsproblem (b) aus § 2.2. Dann ist (5.8) also mit (5.11), (5.12) äquivalent, d.h.

$$\int_0^T \int_{\Omega_1} (Cy - z_d) v^* dx dt = \int_0^T \int_{\Omega_1} |Cy - z_d| dx dt, \quad \|v^*\|_{\infty} = 1.$$

Wegen $(Cy - z_d) v^* \leq |Cy - z_d|$ ist dies gleichbedeutend mit

$$(5.16) \quad (Cy - z_d) v^* = |Cy - z_d|, \quad \|v^*\|_{\infty} = 1.$$

Falls garantiert werden kann, daß $Cy - z_d$ höchstens auf einer Nullmenge verschwindet, erhält man für die optimale Steuerung v^* des Dualproblems ebenfalls ein Bang-Bang-Prinzip:

$$(5.17) \quad |v^*| = 1.$$

II. APPROXIMATION PARABOLISCHER KONTROLLPROBLEME

§6. Approximation konvexer Optimierungsprobleme

6.1. Konvergenzsätze

Wir gehen noch einmal in die allgemeine Situation von §3 zurück. Dabei nehmen wir an, daß X_1 bzw. Y_2 Banach-Räume seien, deren Normen wir mit $\|\cdot\|_1$ bzw.

$\|\cdot\|_2$ bezeichnen. Oberdies mögen eine überall endliche, stetige, konvexe Funktion f und eine überall endliche, stetige, konkave Funktion f_* existieren, so daß

$$(6.1) \quad f(x) = (F0)(x) \quad \forall x \in \text{dom}(F0), \quad f_*(u^*) = (F^*0)(u^*) \quad \forall u^* \in \text{dom}(F^*0).$$

Wir setzen als Abkürzung

$$(6.2) \quad C := \text{dom}(F0), \quad C_* := \text{dom}(F^*0).$$

Weiter seien "diskretisierte" Zielfunktionen $f_n: X_1 \rightarrow \mathbb{R}$, $f_{*n}: Y_2 \rightarrow \mathbb{R}$ gegeben, die wir auch als stetig und konvex bzw. konkav annehmen. Schließlich seien "diskretisierte" Restriktionsmengen C_n , C_{*n} gegeben, die ebenfalls konvex seien. Damit haben wir die folgenden beiden Paare kontinuierlicher und "diskretisierter" Optimierungsprobleme erhalten.

(P) Minimiere $f(x)$ über alle $x \in C$.

(P_n) Minimiere $f_n(x)$ über alle $x \in C_n$.

(D) Maximiere $f_*(u^*)$ über alle $u^* \in C_*$.

(D_n) Maximiere $f_{*n}(u^*)$ über alle $u^* \in C_{*n}$.

Wir wollen in diesem Paragraphen der Frage nachgehen, wann $\lim_{n \rightarrow \infty} \inf(P_n) = \inf(P)$

gilt. Dazu machen wir für die Primalprobleme die folgenden Voraussetzungen.

(V0) Für alle $n \in \mathbb{N}$ gilt

$$(6.3) \quad -\infty < \inf(P) = \sup(D) < +\infty,$$

$$(6.4) \quad -\infty < \inf(P_n) = \sup(D_n) < +\infty.$$

(V1) Zu jedem $x \in C$ existiert eine Folge $\{x_n\} \subset X_1$ mit $x_n \in C_n$ für alle $n \in \mathbb{N}$ und $x = \lim_{n \rightarrow \infty} x_n$.

(V2) Für alle $x \in X_1$ gilt

$$(6.5) \quad \limsup_{n \rightarrow \infty} f_n(x) \leq f(x).$$

(V3) Zu jedem $x \in X$ existiert eine Konstante c_x , so daß

$$(6.6) \quad f_n(x) \geq c_x \quad \forall n \in \mathbb{N}.$$

Analoge Voraussetzungen formulieren wir für die Dualprobleme.

(V1_{*}) Zu jedem $u^* \in C_*$ existiert eine Folge $\{u_n^*\} \subset Y_2$ mit $u_n^* \in C_{*n}$ für alle $n \in \mathbb{N}$ und $u^* = \lim_{n \rightarrow \infty} u_n^*$.

(V2_{*}) Für alle $u^* \in Y_2$ gilt

$$(6.7) \quad \liminf_{n \rightarrow \infty} f_{*n}(u^*) \geq f_*(u^*).$$

(V3_{*}) Zu jedem $u^* \in Y_2$ existiert eine Konstante c_{u^*} mit

$$(6.8) \quad f_{*n}(u^*) \leq c_{u^*} \quad \forall n \in \mathbb{N}.$$

Dann gilt der folgende Konvergenzsatz.

Satz 6.1. Für die oben eingeführten Minimierungsprobleme (P_n) und (P) gilt unter den Voraussetzungen (V0), (V1) - (V3) und (V1_{*}) - (V3_{*}) die Beziehung

$$(6.9) \quad \liminf_{n \rightarrow \infty} (P_n) = \inf(P).$$

Den Beweis von Satz 6.1 geben wir im nächsten Abschnitt.

Wir wenden uns jetzt der Frage zu, wie man aus (6.9) auf die Konvergenz etwa vorhandener Optimallösungen schließen kann.

(V4) Sei $\{x_n\} \subset X_1$ eine beschränkte Folge mit $x_n \in C_n$ für alle $n \in \mathbb{N}$. Dann gibt es eine Teilfolge $\{x_{n_i}\}$, die im Sinne der schwachen Topologie von

X_1 gegen ein gewisses $z \in C$ konvergiert. Für diese Teilfolge gilt überdies

$$(6.10) \quad f(z) \leq \liminf_{i \rightarrow \infty} f_{n_i}(x_{n_i}).$$

Satz 6.2. Zusätzlich zu den in Satz 6.1 gemachten Voraussetzungen gelte (V4). Weiter habe jedes Problem (P_n) eine Optimallösung x_n und die Folge $\{x_n\}$ sei beschränkt. Dann gibt es eine Teilfolge von $\{x_n\}$, die schwach gegen eine Optimallösung von (P) konvergiert. Hat (P) höchstens eine Optimallösung x_0 , so konvergiert die Folge $\{x_n\}$ selbst gegen x_0 .

Beweis: Bezeichne " $\rightharpoonup$ " schwache Konvergenz. Nach (V4) und Satz 6.1 gibt es eine Teilfolge $\{x_{n_i}\}$ von $\{x_n\}$ und ein $z \in C$ mit

$$f(z) \leq \liminf_{i \rightarrow \infty} f_{n_i}(x_{n_i}) = \inf(P).$$

Also ist z Optimallösung für (P) .

Hat (P) höchstens eine Optimallösung x_0 , so folgt sofort $z = x_0$ und $x_n \rightharpoonup x_0$ für $n \rightarrow \infty$. $\square$

Eine analoge Aussage ist für das Dualproblem möglich.

(V4_{*}) Sei $\{u_n^*\} \subset Y_2$ eine beschränkte Folge mit $u_n^* \in C_{*n}$ für alle $n \in \mathbb{N}$.

Dann gibt es eine Teilfolge $\{u_{n_i}^*\}$, die im Sinne der schwachen Topologie von Y_2 gegen ein gewisses $v^* \in C_*$ konvergiert. Für diese Teilfolge gilt überdies

$$(6.11) \quad f_*(v^*) \geq \limsup_{i \rightarrow \infty} f_{*n_i}(u_{n_i}^*).$$

Satz 6.2.* Zusätzlich zu den in Satz 6.1 gemachten Voraussetzungen gelte (V4_{*}).

Weiter habe jedes Problem (D_n) eine Optimallösung u_n^* und die Folge $\{u_n^*\}$ sei beschränkt. Dann gibt es eine Teilfolge von $\{u_n^*\}$, die schwach gegen eine Optimallösung von (D) konvergiert. Hat (D) höchstens eine Optimallösung u_0^* , so konvergiert die Folge $\{u_n^*\}$ selbst gegen u_0^* .

6.2. Der Beweis von Satz 6.1

Wir benötigen zunächst zwei Hilfssätze.

Lemma 6.1. Sei X ein Banach-Raum, $\{f_n\}_{n \in \mathbb{N}}$ eine Familie konvexer, stetiger, auf ganz X definierter Funktionale, die auf X punktweise gleichmäßig nach oben beschränkt sind. Dann gibt es zu beliebigem $x_0 \in X$ eine Konstante c und eine Umgebung W von x_0 mit

$$(6.12) \quad f_n(x) \leq c \quad \forall x \in W, \quad \forall n \in \mathbb{N}.$$

Beweis: Nach einem Prinzip der gleichmäßigen Beschränktheit (vergl. HIRZEBRUCH-SCHARLAU [16], S. 22) gibt es eine offene Menge U und eine Konstante c_0 mit

$$(6.13) \quad f_n(x) \leq c_0 \quad \forall x \in U, \quad \forall n \in \mathbb{N}.$$

Wähle $x_1 \in U$ und $\varrho > 1$ mit $\varrho x_1 \in U$. Setze

$$V := -x_1 + \frac{1}{\varrho} U.$$

V ist Nullumgebung. Jedes $z \in V$ läßt sich in der Form

$$\begin{aligned} z &= -x_1 + \frac{1}{\varrho} y \\ &= \left(1 - \frac{1}{\varrho}\right) \left(-\left(1 - \frac{1}{\varrho}\right)^{-1} x_1\right) + \frac{1}{\varrho} y \end{aligned}$$

mit $y \in U$ darstellen. Wegen (6.13) und der punktweisen Beschränktheit der f_n folgt hieraus die Existenz einer Konstanten c_1 , so daß

$$(6.14) \quad f_n(z) \leq \left(1 - \frac{1}{\varrho}\right) f_n\left(-\left(1 - \frac{1}{\varrho}\right)^{-1} x_1\right) + \frac{1}{\varrho} f_n(y) \leq c_1 \quad \forall z \in V, \quad \forall n \in \mathbb{N}.$$

Setze weiter

$$w := x_0 + \left(1 - \frac{1}{\varrho}\right) V.$$

Dann läßt sich jedes $z \in W$ in der Form

$$\begin{aligned} z &= x_0 + \left(1 - \frac{1}{\varrho}\right) y \\ &= \frac{1}{\varrho} (\varrho x_0) + \left(1 - \frac{1}{\varrho}\right) y \end{aligned}$$

mit $y \in V$ darstellen. Also folgt aus (6.14) und der punktweisen Beschränktheit der f_n die Existenz einer Konstanten c mit

$$f_n(z) \leq \frac{1}{\varrho} f_n(\varrho x_0) + \left(1 - \frac{1}{\varrho}\right) f_n(y) \leq c \quad \forall x \in W, \quad \forall n \in \mathbb{N}.$$

Lemma 6.2. Sei X ein Banach-Raum, $\{f_n\}_{n \in \mathbb{N}}$ eine Familie konvexer, stetiger, auf ganz X definierter Funktionale, die betragsmäßig punktweise gleichmäßig beschränkt sind. Dann gibt es zu beliebigem $x_0 \in X$ eine Konstante c und eine Kugel $K_{\varrho}(x_0)$ mit Radius $\varrho > 0$ um x_0 , so daß

$$(6.15) \quad f_n(x) - f_n(x_0) \leq c \|x - x_0\| \quad \forall x \in K_{\vartheta}(x_0), \quad \forall n \in \mathbb{N}.$$

Beweis: Vorbemerkung: für ein konvexes Funktional $f: X \rightarrow \mathbb{R}$ gilt bekanntlich

$$x, h \in X, \quad 0 < \alpha \leq \beta \Rightarrow$$

(6.16)

$$\frac{1}{\alpha} [f(x + \alpha h) - f(x)] \leq \frac{1}{\beta} [f(x + \beta h) - f(x)].$$

Nach Lemma 6.1 gibt es eine Konstante c_0 und ein $\vartheta > 0$, so daß

$$f_n(x_0 + z) \leq c_0 \quad \forall z \in K_{\vartheta}(0), \quad \forall n \in \mathbb{N}.$$

Aus (6.16) folgt für $x \in K_{\vartheta}(x_0) \setminus \{x_0\}$

$$\begin{aligned} \frac{f_n(x) - f_n(x_0)}{\|x - x_0\|} &= \frac{f_n\left(x_0 + \|x - x_0\| \frac{x - x_0}{\|x - x_0\|}\right) - f_n(x_0)}{\|x - x_0\|} \\ &\leq \frac{1}{\vartheta} [f_n(x_0 + \vartheta \frac{x - x_0}{\|x - x_0\|}) - f_n(x_0)]. \end{aligned}$$

Hieraus folgt mit $c_1 := \sup_{n \in \mathbb{N}} |f_n(x_0)|$ die Behauptung:

$$\frac{f_n(x) - f_n(x_0)}{\|x - x_0\|} \leq \frac{c_0 + c_1}{\vartheta} \quad \forall x \in K_{\vartheta}(x_0) \setminus \{x_0\}, \quad \forall n \in \mathbb{N}.$$

Beweis von Satz 6.1. Sei $\varepsilon > 0$ gegeben. Es gibt ein $x_0 \in C$ mit

$$f(x_0) \leq \inf(P) + \frac{\varepsilon}{3}.$$

Nach den Voraussetzungen (V2), (V3) und Lemma 6.2 gibt es Konstanten $\vartheta > 0$, $c > 0$ mit

$$f_n(x) - f_n(x_0) \leq c \|x - x_0\|_1 \quad \forall x \in K_{\vartheta}(x_0), \quad \forall n \in \mathbb{N}.$$

Setze $\vartheta_0 := \min\{\vartheta, \frac{\varepsilon}{3c}\}$. Nach Voraussetzung (V2) gibt es also ein n_0 , so daß für alle $n \geq n_0$, $x \in K_{\vartheta_0}(x_0)$

$$\begin{aligned} f_n(x) - f(x_0) &= f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0) \\ &\leq c \|x - x_0\|_1 + f_n(x_0) - f(x_0) \leq \frac{2}{3} \varepsilon. \end{aligned}$$

Nach Voraussetzung (V1) gibt es ein n_1 , so daß zu jedem $n \geq n_1$ ein $x_n \in C_n \cap K_{\mathcal{D}_0}(x_0)$ existiert. Also folgt für alle $n \geq n_2 := \max \{n_0, n_1\}$

$$(6.17) \quad \inf(P_n) \leq f_n(x_n) \leq f(x_0) + \frac{2}{3} \varepsilon \leq \inf(P) + \varepsilon.$$

Eine analoge Überlegung führen wir für das Dualproblem durch. Es gibt ein $u_0^* \in C_*$ mit

$$\sup(D) - \frac{\varepsilon}{3} \leq f_*(u_0^*).$$

Nach den Voraussetzungen (V2_{*}) und (V3_{*}) und Lemma 6.2 (angewandt auf die konvexen Funktionen $-f_{*n}$) gibt es Konstanten $\hat{\mathcal{D}}_1 > 0$, $c_1 > 0$, so daß

$$f_{*n}(u_0^*) - f_{*n}(u^*) \leq c_1 \|u^* - u_0^*\|_2 \quad \forall u^* \in K_{\hat{\mathcal{D}}_1}(u_0^*), \quad \forall n \in \mathbb{N}.$$

Setze $\mathcal{D}_1 := \min \left\{ \hat{\mathcal{D}}_1, \frac{\varepsilon}{3c_1} \right\}$. Nach Voraussetzung (V2_{*}) gibt es also ein n_3 , so daß für alle $n \geq n_3$, $u^* \in K_{\mathcal{D}_1}(u_0^*)$

$$\begin{aligned} f_*(u_0^*) - f_{*n}(u^*) &= f_*(u_0^*) - f_{*n}(u_0^*) + f_{*n}(u_0^*) - f_{*n}(u^*) \\ &\leq f_*(u_0^*) - f_{*n}(u_0^*) + c_1 \|u^* - u_0^*\|_2 \leq \frac{2}{3} \varepsilon. \end{aligned}$$

Nach Voraussetzung (V1_{*}) gibt es ein n_4 , so daß zu jedem $n \geq n_4$ ein

$u_n^* \in C_{*n} \cap K_{\mathcal{D}_1}(u_0^*)$ existiert. Also folgt für alle $n \geq n_5 := \max \{n_3, n_4\}$

$$\sup(D) - \varepsilon \leq f_*(u_0^*) - \frac{2}{3} \varepsilon \leq f_{*n}(u_n^*) \leq \sup(D_n).$$

Wegen Voraussetzung (V0) gilt also für alle $n \geq n_5$

$$\inf(P) - \varepsilon \leq \inf(P_n).$$

Zusammen mit (6.17) folgt hieraus die Behauptung.]

6.3. Ein Hilfssatz

Zum Nachweis der Voraussetzung (V2_{*}) werden wir später den folgenden Hilfssatz benötigen.

Lemma 6.3. Seien C_n und C konvexe Teilmengen eines Banach-Raumes X und sei $f: X \rightarrow \mathbb{R}$ ein konvexes, auf beschränkten Mengen gleichmäßig stetiges Funktional. Falls C nicht beschränkt ist, konvergiere $\{f(x_n)\}$ gegen ∞ , wenn

$\{x_n\}$ gegen ∞ konvergiert. Weiter existiere in diesem Fall zu jedem $n \in \mathbb{N}$ ein $x_n^0 \in C_n$, so daß $f(x_n^0)$ nach oben beschränkt ist. Schließlich existiere zu jeder Folge $\{x_n\}$ mit $x_n \in C_n$ für alle $n \in \mathbb{N}$ eine Folge $\{\tilde{x}_n\} \subset C$, so daß $\lim_{n \rightarrow \infty} \|x_n - \tilde{x}_n\| = 0$. Dann gilt

$$(6.18) \quad \inf_{x \in C} f(x) \leq \liminf_{n \rightarrow \infty} \left(\inf_{x \in C_n} f(x) \right).$$

Beweis: Sei $\varepsilon > 0$ gegeben. Zu jedem $n \in \mathbb{N}$ gibt es ein $x_n \in C_n$ mit

$$(6.19) \quad f(x_n) \leq \inf_{x \in C_n} f(x) + \frac{\varepsilon}{2}.$$

Zu $\{x_n\}$ existiert eine Folge $\{\tilde{x}_n\} \subset C$ mit $\lim_{n \rightarrow \infty} \|x_n - \tilde{x}_n\| = 0$. Nach unseren

Voraussetzungen liegen diese beiden Folgen in einer beschränkten Menge. Wäre etwa $\{x_n\}$ nicht beschränkt, so gäbe es eine gegen ∞ strebende Teilfolge $\{x_{n_i}\}$ und aus (6.19) würde $\lim_{i \rightarrow \infty} \left(\inf_{x \in C_{n_i}} f(x) \right) = \infty$ folgen, was wegen $\inf_{x \in C_{n_i}} f(x) \leq f(x_{n_i}^0)$

nicht sein kann. Es gibt daher ein n_0 , so daß

$$|f(x_n) - f(\tilde{x}_n)| \leq \frac{\varepsilon}{2} \quad \forall n \geq n_0.$$

Hieraus folgt mit (6.19) für alle $n \geq n_0$

$$\inf_{x \in C} f(x) \leq f(\tilde{x}_n) \leq \inf_{x \in C_n} f(x) + \varepsilon,$$

also die Behauptung. $\square$

§7. Diskretisierungen parabolischer Kontrollprobleme

7.1. Das diskrete primale Problem

Die Diskretisierung des Kontrollproblems vollziehen wir in zwei Schritten. Erst diskretisieren wir die Menge U_{ad} der zulässigen Steuerungen und dann die Differentialgleichung. Wir nehmen an, daß es nicht erforderlich sei, die Zielfunktion zu diskretisieren, um das Kontrollproblem numerisch behandeln zu können.

Wir beginnen also mit der Diskretisierung von U_{ad} . Die Dimension eines beliebigen Vektorraumes V bezeichnen wir mit $\dim V$. Seien zu jedem $l \in \mathbb{N}$ ein Vektorraum U_l mit

$$(7.1) \quad U_l \subset U, \quad \dim U_l < \infty$$

und nichtleere, abgeschlossene und konvexe Mengen

$$(7.2) \quad U_{ad}^l \subset U_l$$

gegeben. Natürlich benötigen wir Voraussetzungen, die garantieren, daß die Mengen U_{ad}^l die Menge U_{ad} in geeigneter Weise approximieren. Wir kommen darauf im nächsten Paragraphen zurück.

Um die Differentialgleichung zu diskretisieren, stellen wir zunächst fest, daß die Gleichungen (2.4) und (2.5) mit

$$(7.3) \quad \left[z, \frac{dy}{dt} \right] + \int_0^T a(t, z(t), y(t)) dt + (z(0), y(0) - y_0) = [z, e + Bu] \quad \forall z \in W(0, T)$$

äquivalent sind. Man sieht dies sofort ein, wenn man beachtet, daß der Raum

$$(7.4) \quad \{ z \in W(0, T) \mid z(0) = 0 \}$$

dicht in $L^2(0, T; V)$ liegt (LIONS - MAGENES [31], Kap III), und daß $z(0)$ den Raum H durchläuft, wenn z den Raum $W(0, T)$ durchläuft (vergl. § 1.4). Wir nehmen an, daß zu allen $m, n \in \mathbb{N}$ Vektorräume S_n, V_m mit

$$(7.5) \quad S_n \subset H^1(]0, T[), \quad \dim S_n < \infty,$$

$$(7.6) \quad V_m \subset V, \quad \dim V_m < \infty$$

gegeben seien.

Für eine beliebige Teilmenge M eines Vektorraumes V bezeichne $\dim M$ die

lineare Hülle von M in V . Dann sei

$$(7.7) \quad W_{mn} := \text{lin} \{ sw \mid s \in S_n, w \in V_m \}.$$

Wir diskretisieren die Differentialgleichung (2.4), (2.5), indem wir (7.3) nur für alle $z \in W_{mn}$ fordern, also

$$(7.8) \quad [z, \frac{dy}{dt}] + \int_0^1 a(t, z(t), y(t)) dt + (z(0), y(0) - y_0) = [z, e + Bu] \quad \forall z \in W_{mn}.$$

Die Lösung von (7.8) suchen wir in W_{mn} . Im nächsten Abschnitt und in §9 wird sich zeigen, daß diese Konstruktion tatsächlich sinnvoll ist.

Als diskretes Kontrollproblem führen wir nun ein:

$$(P_{mn}^1) \quad \text{Minimiere } f(Cy - z_d) + \langle u, Nu \rangle_U \\ \text{über alle } y \in W_{mn} \text{ und } u \in U_{ad}^1 \text{ unter den Nebenbedingungen (7.8).}$$

Das diskrete Problem (P_{mn}^1) ist in dem Sinne endlichdimensional, als daß die zugrunde liegenden Funktionenräume endlichdimensional sind. Um das Problem mit einem numerischen Verfahren behandeln zu können, muß man in den Funktionenräumen Basen einführen und das Problem so umschreiben, daß die Koeffizienten der Funktionen bezüglich dieser Basen gesucht werden. Dies ist im allgemeinen natürlich nicht ohne einen weiteren Approximationsprozeß möglich (z.B. weil man gewisse Integrationen nicht exakt durchführen kann oder weil die Auswertung der Zielfunktion nicht exakt möglich ist). In vielen Beispielen kann man dagegen diese Rechenoperation exakt durchführen, so daß wir (P_{mn}^1) als "vollständig diskretisierte" Version des Problems (P) ansehen wollen. Auf den Übergang zu den Koeffizienten kommen wir später noch einmal zurück, insbesondere in § 11.

7.2. Das Dualproblem des diskretisierten Kontrollproblems

Um Satz 6.1 anwenden zu können, stellen wir zunächst zu dem Problem (P_{mn}^1) ein Dualproblem auf.

Für den Beweis des folgenden Lemmas 7.1 und für den in §9 zu führenden Konvergenzbeweis ist es von Vorteil, wenn Voraussetzung (1.34) mit $\lambda = 0$ erfüllt werden kann. Wir zeigen, daß dies o.B.d.A. möglich ist. y ist Lösung von (2.4), (2.5) genau dann, wenn $z := e^{-\lambda t} y$ Lösung von

$$(7.9) \quad \frac{dz}{dt} + (A(t) + \lambda I)z = e^{-\lambda t} e e^{-\lambda t} Bu, \quad z(0) = y_0$$

ist (vergl. LIONS [31], Kap. III). Hierbei sei I die Identität von $L^2(0, T; V)$. Der Operator $A(t) + \lambda I$ wird von

$$(7.10) \quad \tilde{a}(t, u, v) := a(t, u, v) + \lambda(u, v)$$

erzeugt (im Sinne von (1.35)). $\tilde{a}$ erfüllt aber Voraussetzung (1.34) mit $\lambda = 0$. Wir ersetzen jetzt in (P) a durch $\tilde{a}$, e durch $e^{-\lambda t} e$, B durch $e^{-\lambda t} B$ und C durch $Ce^{\lambda t}$ (d.h. $(Ce^{\lambda t})(z) = C(e^{\lambda t} z)$). Dann ergibt sich ein neues Problem $(\tilde{P})$, das aber wieder die Struktur von (P) hat. Ist das Paar (z, u) Optimallösung von $(\tilde{P})$, so ist $(e^{\lambda t} z, u)$ Optimallösung von (P) und es gilt $\inf(P) = \inf(\tilde{P})$. Wir nehmen ab jetzt an, daß $\lambda = 0$ gilt.

Sei W_{mn} mit dem von $W(0, T)$ induzierten Skalarprodukt versehen. Dann ist W_{mn} ein endlichdimensionaler Hilbert-Raum. Für $z, y \in W_{mn}$ setzen wir

$$(7.11) \quad b(z, y) := \left[z, \frac{dy}{dt} \right] + \int_0^T a(t, z(t), y(t)) dt + (z(0), y(0)).$$

Für $y \in W_{mn}$ wird durch $\tilde{D}_{mn}(y)(z) := b(z, y)$, $z \in W_{mn}$, ein lineares Funktional auf W_{mn} definiert. Da W_{mn} endlichdimensional ist, ist dieses auch stetig, so daß $\tilde{D}_{mn}$ ein linearer Operator von W_{mn} nach W_{mn}^* ist. Da W_{mn} ein Hilbert-Raum ist, gibt es einen stetigen Isomorphismus $j_1: W_{mn}^* \rightarrow W_{mn}$ mit stetiger inverser Abbildung, so daß $\langle j_1(w^*), w \rangle_{W(0, T)} = w^*(w)$ für alle $w \in W_{mn}$, $w^* \in W_{mn}^*$ gilt. Wir setzen $D_{mn} := j_1 \circ \tilde{D}_{mn}$ und haben dann also

$$(7.12) \quad \langle z, D_{mn} y \rangle_{W(0, T)} = b(z, y) \quad \forall z, y \in W_{mn}.$$

Lemma 7.1. Der oben konstruierte Operator D_{mn} ist ein stetiger Isomorphismus von W_{mn} auf W_{mn} mit stetiger inverser Abbildung.

Beweis: Wäre D_{mn} nicht injektiv, so gäbe es ein $y \in W_{mn} \setminus \{0\}$ mit

$$(7.13) \quad b(y, y) = 0.$$

Für alle $y \in W(0, T)$ gilt bekanntlich die Formel

$$(7.14) \quad \left[y, \frac{dy}{dt} \right] = \frac{1}{2} |y(T)|^2 - \frac{1}{2} |y(0)|^2$$

(vergl. LIONS-MAGENES [31], Kap. I). Zusammen mit (1.34) folgt hieraus

$$\frac{1}{2}|y(T)|^2 + \frac{1}{2}|y(0)|^2 + \alpha \int_0^T \|y(t)\|^2 dt \leq 0,$$

also $y = 0$ im Widerspruch zu unserer Annahme.

Da D_{mn} also injektiv ist, folgt $\dim D_{mn}(W_{mn}) = \dim W_{mn}$. Daher ist D_{mn} surjektiv. Die Stetigkeit von D_{mn} und D_{mn}^{-1} folgt aus der Tatsache, daß W_{mn} endlichdimensional ist. $\square$

Wir definieren einen linearen Operator $j_2: L^2(0, T; V^*) \rightarrow W_{mn}^*$ durch

$$(7.15) \quad j_2(u)(y) := [y, u] \quad \forall y \in W_{mn}, \forall u \in L^2(0, T; V^*).$$

Da W_{mn} endlichdimensional ist, sind die von $W(0, T)$ und $L^2(0, T; V)$ induzierten Normen äquivalent. Es gibt also ein $c > 0$ mit

$$\|y\|_{L^2(0, T; V)} \leq c \|y\|_{W(0, T)} \quad \forall y \in W_{mn}.$$

Konvergiere $\{u_n\}$ gegen u in $L^2(0, T; V^*)$. Dann folgt für alle $y \in W_{mn}$

$$|j_2(u_n)(y) - j_2(u)(y)| = |[y, u_n - u]| \\ \leq c \|y\|_{W(0, T)} \|u_n - u\|_{L^2(0, T; V^*)}.$$

Also konvergiert $\{j_2(u_n)\}$ gegen $j_2(u)$. Folglich ist j_2 stetig.

Sei $j := j_1 \circ j_2$. Dann gilt für alle $y \in W_{mn}$, $u \in L^2(0, T; V^*)$

$$(7.16) \quad \langle y, j(u) \rangle_{W(0, T)} = [y, u].$$

Weiter gibt es ein $\tilde{y}_0 \in W_{mn}$, so daß

$$\langle y, \tilde{y}_0 \rangle_{W(0, T)} = (y(0), y_0) \quad \forall y \in W_{mn}.$$

Also ist (7.8) äquivalent mit

$$(7.17) \quad D_{mn}y = \tilde{y}_0 + j(e) + j(Bu).$$

Zusammen mit Lemma 7.1 ergibt sich hieraus der folgende Satz.

Satz 7.1. Falls Voraussetzung (1.34) mit $\lambda = 0$ erfüllt werden kann, ist Gleichung (7.8) für alle $u \in U$ durch genau ein $y \in W_{mn}$ lösbar.

Mit diesen Vorbereitungen sind wir in der Lage, die Theorie aus §3 auf das diskretisierte Kontrollproblem (P_{mn}^1) anzuwenden. Wir setzen dazu

$$(7.18) \quad \begin{aligned} X_1 &:= W_{mn} \oplus U_1, & Y_1 &:= W_{mn} \oplus U_1, \\ X_2 &:= W_{mn} \oplus B, & Y_2 &:= W_{mn} \oplus B^*. \end{aligned}$$

Weiter sei für alle $w \in W_{mn}$

$$(7.19) \quad K_w^d := \{(y, u) \in X_1 \mid D_{mn}y = -w + \tilde{y}_0 + j(e) + (jB)(u), u \in U_{ad}^1\}.$$

Man kann genau wie in §4.1 eine Bifunktion F definieren, nur mit K_w^d anstelle von K_w . Mit analogen Überlegungen wie im kontinuierlichen Fall ergibt sich dann ein Dualproblem und ein Dualitätssatz. Wir führen diese Überlegungen nicht im Detail durch, sondern teilen nur das Ergebnis mit.

Als Dualproblem zu (P_{mn}^1) erhält man:

(D_{mn}^1) Maximiere

$$(7.20) \quad -f^*(v^*) + \min_{u \in U_{ad}^1} \langle u, Nu + B^*w^* \rangle_U - \langle z_d, v^* \rangle_B + [w^*, e] + (y_0, w^*(0))$$

über alle $w^* \in W_{mn}$ und $v^* \in B^*$ unter der Nebenbedingung

$$(7.21) \quad -[z, \frac{dw^*}{dt}] + \int_0^T a(t, w^*(t), z(t)) dt + (z(T), w^*(T)) = [z, C^*v^*] \quad \forall z \in W_{mn}.$$

Satz 7.2. Sei Voraussetzung (2.2) für U_{ad}^1 anstelle von U_{ad} erfüllt. Dann hat sowohl (P_{mn}^1) als auch (D_{mn}^1) eine Optimallösung und es gilt

$$(7.22) \quad \min(P_{mn}^1) = \max(D_{mn}^1).$$

§8. Approximation der Menge der zulässigen Steuerungen

8.1. Allgemeine Voraussetzungen

Es erscheint zunächst naheliegend, die Mengen U_{ad}^1 aus §7.1 durch

$$(8.1) \quad U_{ad}^1 := U_{ad} \cap U_1$$

zu definieren. Andererseits ist es für die numerische Behandlung von (P_{mn}^1) erforderlich, die durch

$$(8.2) \quad \tilde{U}_{ad}^1 := \left\{ \lambda \in \mathbb{R}^{1d} \mid \sum_{i=1}^{1d} \lambda_i u_i \in U_{ad}^1 \right\}$$

definierte Menge explizit zu kennen. Hierbei sei $1_d := \dim U_1$ und $\{u_1, \dots, u_{1_d}\}$ eine Basis von U_1 . Wird U_{ad}^1 gemäß (8.1) gewählt, so läßt sich $\tilde{U}_{ad}^1$ im allgemeinen nicht explizit bestimmen. Man ist daher gezwungen, auch Fälle zuzulassen, in denen nicht $U_{ad}^1 \subset U_{ad}$ gilt. Wir werden hierfür sogleich ein Beispiel geben.

Wir führen die folgenden Voraussetzungen ein.

$$(8.3) \quad \begin{aligned} &\text{Zu jedem } u \in U_{ad} \text{ gibt es eine Folge } \{u_l\} \subset U \text{ mit } u_l \in U_{ad}^1 \text{ für alle} \\ &l \in \mathbb{N} \text{ und } u = \lim_{l \rightarrow \infty} u_l. \end{aligned}$$

$$(8.4) \quad \begin{aligned} &\text{Zu jeder Folge } \{u_l\} \subset U \text{ mit } u_l \in U_{ad}^1 \text{ für alle } l \in \mathbb{N} \text{ existiert eine} \\ &\text{Folge } \{\tilde{u}_l\} \subset U \text{ mit } \tilde{u}_l \in U_{ad} \text{ für alle } l \in \mathbb{N} \text{ und } \lim_{l \rightarrow \infty} \|u_l - \tilde{u}_l\|_U = 0. \end{aligned}$$

Für den Nachweis von (8.3) ist in vielen Fällen das folgende Lemma nützlich.

Lemma 8.1. Es existiere eine in U_{ad} dicht liegende Teilmenge $V_{ad} \subset U_{ad}$, so daß zu jedem $v \in V_{ad}$ eine Folge $\{v_l\} \subset U$ mit $v_l \in U_{ad}^1$ für alle $l \in \mathbb{N}$ und $v = \lim_{l \rightarrow \infty} v_l$ existiert. Dann ist Voraussetzung (8.3) erfüllt.

Beweis: Sei $u \in U_{ad}$ gegeben. Zu jedem $k \in \mathbb{N}$ gibt es ein $v_k \in V_{ad}$ mit $\|u - v_k\|_U < \frac{1}{k}$ und eine Folge $\{v_k^\mu\}_{\mu \in \mathbb{N}}$ mit $v_k = \lim_{\mu \rightarrow \infty} v_k^\mu$ und $v_k^\mu \in U_{ad}^1$ für alle $\mu \in \mathbb{N}$. Also existiert zu jedem $k \in \mathbb{N}$ ein μ_k mit

$$\|v_k - v_k^{\mu_k}\|_U < \frac{1}{k} \quad \forall \mu \geq \mu_k.$$

Hierbei können wir o.B.d.A. annehmen, daß die Folge $\{\mu_k\}$ streng monoton gegen ∞ strebt. Zu jedem $l \in \mathbb{N}$, $l \geq \mu_1$ gibt es daher eindeutig $k, \nu \in \mathbb{N}$ mit $0 \leq \nu \leq \mu_{k+1} - \mu_k - 1$ und $l = \mu_k + \nu$. Wir setzen $u_l := v_k^{\mu_k + \nu}$ und definieren eine Abbildung $j: \mathbb{N} \setminus \{1, \dots, \mu_1 - 1\} \rightarrow \mathbb{N}$ durch $j(l) := k$. Dann gilt

$$\|v_{j(l)} - u_l\|_U < \frac{1}{j(l)}, \quad u_l \in U_{ad} \quad \forall l \geq \mu_1.$$

Also folgt

$$\|u - u_l\|_U \leq \|u - v_{j(l)}\|_U + \|v_{j(l)} - u_l\|_U < \frac{2}{j(l)}.$$

Da $\{\frac{2}{j(l)}\}$ eine monoton fallende Nullfolge ist, haben wir hiermit die Behauptung bewiesen.]

8.2 Beispiele

Die Voraussetzungen (8.3), (8.4) werden wir in §10 benötigen, um Satz 6.1 anzuwenden. Wir wollen sie zunächst im Falle von zwei einfachen Beispielen nachweisen.

1. Sei $m > 0$, $Q := [0, T] \times [b, c]$, $U := L^2(Q)$ und $U_{ad} := \{u \in L^2(Q) \mid |u| \leq m\}$.

Zu allen $l_1, l_2 \in \mathbb{N}$ seien Zerlegungen

$$(8.5) \quad 0 = t_0 < t_1 < \dots < t_{l_1} = T, \quad b = x_0 < x_1 < \dots < x_{l_2} = c,$$

gegeben. Für $i=1, \dots, l_1$; $j=1, \dots, l_2$ sei

$$(8.6) \quad u_{ij}(t, x) := \begin{cases} 1, & \text{falls } (t, x) \in [t_{i-1}, t_i] \times [x_{j-1}, x_j], \\ 0, & \text{sonst.} \end{cases}$$

Setze $U_{l_1, l_2} := \text{lin}\{u_{ij} \mid i=1, \dots, l_1; j=1, \dots, l_2\}$ und wähle $U_{ad}^{l_1, l_2}$ gemäß (8.1).

Dann gilt

$$(8.7) \quad U_{ad}^{l_1, l_2} = \{\lambda \in \mathbb{R}^{l_1 l_2} \mid |\lambda_{ij}| \leq m, i=1, \dots, l_1; j=1, \dots, l_2\}.$$

Lemma 8.2. Falls die Feinheiten der Zerlegungen (8.5) gegen Null streben,

wenn l_1 und l_2 gegen ∞ konvergieren, so erfüllen die Mengen $U_{ad}^{l_1, l_2}$, U_{ad} die Voraussetzungen (8.3), (8.4).

Beweis: Wir wenden Lemma 8.1 an.

Da $C(Q)$ dicht in $L^2(Q)$ liegt, gibt es zu beliebigem $u \in U_{ad}$ und $\varepsilon > 0$ ein $\tilde{v} \in C(Q)$ mit $\|u - \tilde{v}\|_{L^2(Q)} < \varepsilon$. Wir definieren

$$(8.8) \quad v(t, x) := \begin{cases} \tilde{v}(t, x) & , \text{ falls } |\tilde{v}(t, x)| \leq m, \\ m \operatorname{sign} \tilde{v}(t, x), & \text{ falls } |\tilde{v}(t, x)| > m. \end{cases}$$

v ist stetig und für fast alle $(t, x) \in Q$ gilt

$$(8.9) \quad |u(t, x) - v(t, x)| \leq |u(t, x) - \tilde{v}(t, x)|.$$

Also haben wir $v \in U_{ad}$ und $\|u - v\|_{L^2(Q)} < \varepsilon$. Folglich liegt $U_{ad} \cap C(Q)$ dicht in U_{ad} .

Sei ein beliebiges $v \in U_{ad} \cap C(Q)$ gegeben. Für $i=1, \dots, l_1$, $j=1, \dots, l_2$ setze

$$(8.10) \quad v_{1, l_2} := \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} v(t_i, x_j) u_{ij}.$$

Es gilt $v_{1, l_2} \in U_{ad}^{1, l_2}$. Da v gleichmäßig stetig ist, gibt es zu

$\varepsilon_1 := \left(\frac{\varepsilon}{T(c-b)} \right)^{\frac{1}{2}}$ ein $\delta > 0$, so daß $|v(t, x) - v(\bar{t}, \bar{x})| < \varepsilon_1$ gilt, falls $|t - \bar{t}| < \delta$, $|x - \bar{x}| < \delta$. Daher folgt

$$(8.11) \quad \int_0^T \int_b^c |v(t, x) - v_{1, l_2}(t, x)|^2 dx dt = \\ = \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \int_{t_{i-1}}^{t_i} \int_{x_{j-1}}^{x_j} |v(t, x) - v(t_i, x_j)|^2 dx dt < \varepsilon,$$

falls die Feinheiten der Zerlegungen kleiner als δ sind. Also konvergiert $\{v_{1, l_2}\}$ im Sinne der Norm von $L^2(Q)$ gegen v . $\square$

2. Sei $U := L^2[a, b]$ und sei $\varphi \in C[a, b]$ mit $\varphi(x) > 0$ für alle $x \in [a, b]$ gegeben. Setze

$$(8.12) \quad U_{ad} := \{u \in L^2[a, b] \mid |u| \leq \varphi\}.$$

Weiter sei zu jedem $l \in \mathbb{N}$ eine Zerlegung

$$(8.13) \quad a = x_0 < x_1 < \dots < x_l = b$$

von $[a, b]$ gegeben. Wir führen die folgenden Intervalle und Funktionen ein.

$$(8.14) \quad I_i := [x_{i-1}, x_i], \quad i=1, \dots, l,$$

$$g_i^-(x) := \frac{x - x_{i-1}}{x_i - x_{i-1}}, \quad i=1, \dots, l,$$

$$g_i^+(x) := \frac{x - x_{i+1}}{x_i - x_{i+1}}, \quad i=0, \dots, l-1,$$

$$(8.15) \quad u_i(x) := \begin{cases} g_i^-(x), & \text{falls } x \in I_i, \\ g_i^+(x), & \text{falls } x \in I_{i+1}, \quad (i=1, \dots, l-1) \\ 0, & \text{sonst,} \end{cases}$$

$$(8.16) \quad u_0(x) := \begin{cases} g_0^+(x), & \text{falls } x \in I_1, \\ 0, & \text{sonst,} \end{cases}$$

$$(8.17) \quad u_l(x) := \begin{cases} g_l^-(x), & \text{falls } x \in I_l, \\ 0, & \text{sonst.} \end{cases}$$

Sei $U_1 := \text{lin} \{u_0, \dots, u_l\}$. Setze

$$(8.18) \quad \varphi_1 := \sum_{i=0}^l \varphi(x_i) u_i,$$

$$(8.19) \quad U_{ad}^1 := \{u \in U_1 \mid |u| \leq \varphi_1\}.$$

Man sieht sofort, daß

$$(8.20) \quad \tilde{U}_{ad}^1 = \{ \lambda \in \mathbb{R}^{l+1} \mid |\lambda_i| \leq \varphi(x_i), i=0, \dots, l \}.$$

Hier gilt im allgemeinen nicht $U_{ad}^1 \subset U_{ad}$.

Lemma 8.3. Falls die Feinheiten der Zerlegungen (8.13) gegen Null streben, wenn l gegen ∞ konvergiert, so erfüllen die Mengen U_{ad}^1, U_{ad} die Voraussetzungen (8.3), (8.4).

Beweis: (a) Analog wie im Beweis von Lemma 8.2 können wir zeigen, daß $U_{ad} \cap C[a, b]$ dicht in U_{ad} liegt.

Zu beliebigem $v \in U_{ad} \cap C[a, b]$ definieren wir

$$v_1 := \sum_{i=0}^l v(x_i) u_i.$$

Auf allen $[x_{i-1}, x_i]$ gilt

$$v_1(x) = \frac{v(x_i) - v(x_{i-1})}{x_i - x_{i-1}} (x - x_{i-1}) + v(x_{i-1}).$$

Hieraus folgt für alle $x \in [x_{i-1}, x_i]$ die Ungleichung

$$|v(x) - v_1(x)| \leq 2 \sup_{y, z \in I_i} |v(y) - v(z)|$$

und mit einer zu (8.11) analogen Abschätzung zeigt man (8.3).

(b) Sei eine Folge $\{u_l\}$ mit $u_l \in U_{ad}^1$ für alle $l \in \mathbb{N}$ gegeben. Setze

$$(8.21) \quad \tilde{u}_1(x) := \begin{cases} u_1(x), & \text{falls } |u_1(x)| \leq \varphi(x), \\ \varphi(x) \operatorname{sign} u_1(x), & \text{falls } |u_1(x)| > \varphi(x). \end{cases}$$

Dann gilt $\tilde{u}_1 \in U_{ad}$ und

$$(8.22) \quad |u_1(x) - \tilde{u}_1(x)| \leq |\rho_1(x) - \rho(x)| \quad \forall x \in [a, b].$$

(8.22) sieht man sofort ein, wenn man berücksichtigt, daß

$$|u_1(x)| \leq \rho_1(x)$$

für alle $x \in [a, b]$ gilt. Aus (8.22) folgt

$$\|u_1 - \tilde{u}_1\|_{L^2[a, b]} \leq \|\rho_1 - \rho\|_{L^2[a, b]}.$$

Da $\{\rho_1\}$ im Sinne von $L^2[a, b]$ gegen ρ strebt, ist (8.4) nachgewiesen. $\square$

§9. Galerkin-Approximationen parabolischer Differentialgleichungen

Nach dem Konzept aus §7 wollen wir die Differentialgleichung (1.37), (1.38) durch Näherungslösungen $y_{mn} \in W_{mn}$ approximieren, die sich aus

$$(9.1) \quad [z, \frac{dy_{mn}}{dt}] + \int_0^T a(t, z(t), y_{mn}(t)) dt + (z(0), y_{mn}(0) - y_0) = [z, e] \quad \forall z \in W_{mn}$$

ergeben. Wir haben in §7 bereits gezeigt, daß (9.1) eine eindeutig bestimmte Lösung y_{mn} in W_{mn} hat.

Für die approximierenden Teilräume V_m, S_n machen wir die folgenden Voraussetzungen.

(9.2) Zu jedem $w \in V$ gibt es eine Folge $\{w_m\} \subset V$ mit $w_m \in V_m$ für alle $m \in \mathbb{N}$

$$\text{und } w = \lim_{m \rightarrow \infty} w_m.$$

(9.3) Zu jedem $s \in H^1([0, T])$ gibt es eine Folge $\{s_n\} \subset H^1([0, T])$ mit

$$s_n \in S_n \text{ für alle } n \in \mathbb{N} \text{ und } s = \lim_{n \rightarrow \infty} s_n.$$

Satz 9.1. Unter den Voraussetzungen (9.2) und (9.3) gelten die folgenden Aussagen.

a) Sei $\{e_k\}$ eine beschränkte Folge in $L^2(0, T; V^*)$ und y_{mn}^k die zur rechten Seite e_k gebildete Lösung von (9.1). Dann gibt es eine Teilfolge von $\{y_{mn}^k, e_k\}$, die in der schwachen Topologie von $L^2(0, T; V) \oplus L^2(0, T; V^*)$ gegen ein gewisses (y, e) konvergiert, das der Differentialgleichung (1.37), (1.38) genügt.

b) Sei $e \in L^2(0, T; V^*)$ gegeben und sei y_{mn} die Lösung von (9.1). Dann konvergiert $\{y_{mn}\}$ in der Normtopologie von $L^2(0, T; V)$ gegen die Lösung y der Differentialgleichung (1.37), (1.38) und $\{y_{mn}(0)\}$ konvergiert in der Normtopologie von H gegen y_0 .

Beweis: Wir führen den Beweis mit einem ähnlichen Kompaktheitsschluß, wie er von LIONS beim Beweis von Satz 1.2 in LIONS [31], Kapitel III, angewendet wird.

Aus (9.1) folgt

$$(9.4) \quad \left[y_{mn}^k, \frac{dy_{mn}^k}{dt} \right] + \int_0^T a(t, y_{mn}^k(t), y_{mn}^k(t)) dt + (y_{mn}^k(0), y_{mn}^k(0) - y_0) = [y_{mn}^k, e_k].$$

Hieraus erhalten wir mit (7.14) und (1.34) (beachte $\lambda = 0$)

$$\begin{aligned} |y_{mn}^k(T)|^2 + |y_{mn}^k(0)|^2 + 2\alpha \int_0^T \|y_{mn}^k(t)\|^2 dt &\leq 2[y_{mn}^k, e_k] + 2(y_{mn}^k(0), y_0) \\ &\leq 2 \int_0^T \frac{1}{\sqrt{\alpha}} \|y_{mn}^k(t)\| \frac{1}{\sqrt{\alpha}} \|e_k(t)\|_{V^*} dt + 2|y_{mn}^k(0)| |y_0| \\ &\leq \alpha \int_0^T \|y_{mn}^k(t)\|^2 dt + \frac{1}{\alpha} \int_0^T \|e_k(t)\|_{V^*}^2 dt + 2|y_{mn}^k(0)| |y_0|. \end{aligned}$$

Also gilt

$$(9.5) \quad |y_{mn}^k(T)|^2 + (|y_{mn}^k(0)| - |y_0|)^2 + \int_0^T \|y_{mn}^k(t)\|^2 dt \leq \frac{1}{\alpha} \int_0^T \|e_k(t)\|_{V^*}^2 dt + |y_0|^2.$$

Die Folge der $(y_{mn}^k, y_{mn}^k(T), e_k)$ liegt also in einer beschränkten Menge

von $L^2(0, T; V) \oplus H \oplus L^2(0, T; V^*)$. Daher gibt es eine Teilfolge $\{(y_{m_i n_i}^{k_i}, y_{m_i n_i}^{k_i}(T), e_{k_i})\}$,

die in der schwachen Topologie von $L^2(0, T; V) \oplus H \oplus L^2(0, T; V^*)$ gegen ein gewisses (z, z_T, e) aus diesem Raum konvergiert. Aus (9.1) folgt, daß

$$\begin{aligned} (9.6) \quad & \int_0^T (\dot{s}(t)w, y_{mn}^k(t)) dt + \int_0^T a(t, s(t)w, y_{mn}^k(t)) dt + (s(T)w, y_{mn}^k(T)) - (s(0)w, y_0) \\ & = \int_0^T (s(t)w, e_k(t)) dt \end{aligned}$$

für alle $s \in S_n$, $w \in V_m$ gilt. Seien $s \in H^1([0, T])$ und $w \in V$ beliebig gegeben. Aufgrund der Voraussetzungen (9.2), (9.3) gibt es Folgen $\{s_n\}, \{w_m\}$ mit

$$(9.7) \quad s_n \in S_n \quad \forall n \in \mathbb{N}, \quad \lim_{n \rightarrow \infty} \|s_n - s\|_{H^1([0, T])} = 0,$$

$$(9.8) \quad w_m \in W_m \quad \forall m \in \mathbb{N}, \quad \lim_{m \rightarrow \infty} \|w_m - w\| = 0.$$

Damit konvergiert auch $\{s_n\}$ gegen s und $\{\dot{s}_n\}$ gegen $\dot{s}$ in $L^2(0,T)$ und $\{s_n(0)\}$ gegen $s(0)$ und $\{s_n(T)\}$ gegen $s(T)$. Also konvergiert $\{s_n w_m\}$ gegen sw in $L^2(0,T;V)$, denn es gilt

$$\|s_n w_m - sw\|_{L^2(0,T;V)} \leq \|s_n w_m - s w_m\|_{L^2(0,T;V)} + \|s w_m - sw\|_{L^2(0,T;V)}$$

$$= \|w_m\| \|s_n - s\|_{L^2(0,T)} + \|s\|_{L^2(0,T)} \|w_m - w\|.$$

Entsprechend sieht man, daß $\{\dot{s}_n w_m\}$ gegen $\dot{s}w$, $\{s_n(0)w_m\}$ gegen $s(0)w$ und $\{s_n(T)w_m\}$ gegen $s(T)w$ konvergiert.

Zwischenbemerkung: Sei H ein Hilbert-Raum. Konvergiere $\{w_n\} \subset H^*$ in der Normtopologie von H^* gegen w und konvergiere $\{y_n\} \subset H$ in der schwachen Topologie von H gegen y . Weiter existiere ein $M \in \mathbb{R}$ mit $\|y_n\|_H \leq M$ für alle $n \in \mathbb{N}$. Dann konvergiert $\{w_n(y_n)\}$ gegen $w(y)$.

Der Beweis ergibt sich aus der Abschätzung

$$\begin{aligned} |w_n(y_n) - w(y)| &\leq |w_n(y_n) - w(y_n)| + |w(y_n) - w(y)| \\ &\leq M \|w_n - w\|_{H^*} + |w(y_n) - w(y)|. \end{aligned}$$

Aus (9.6) folgt, daß für alle $i \in \mathbb{N}$

$$\begin{aligned} (9.9) \quad & - \int_0^T (\dot{s}_{n_i}(t) w_{m_i}, y_{m_i n_i}^{k_i}(t)) dt + \int_0^T a(t, s_{n_i}(t) w_{m_i}, y_{m_i n_i}^{k_i}(t)) dt \\ & + (s_{n_i}(T) w_{m_i}, y_{m_i n_i}^{k_i}(T)) - (s_{n_i}(0) w_{m_i}, y_0) = \int_0^T (s_{n_i}(t) w_{m_i}, e_{k_i}(t)) dt \end{aligned}$$

gilt. Beachtet man, daß die durch $y \mapsto A^*(\cdot)y(\cdot)$ auf $L^2(0,T;V)$ definierte Abbildung stetig ist (LIONS-MAGENES [31], Kap. III), so kann man also in (9.9.) zum Limes übergehen. Es gilt daher für alle $s \in H^1(0,T)$, $w \in V$

$$\begin{aligned} (9.10) \quad & - \int_0^T (\dot{s}(t)w, z(t)) dt + \int_0^T a(t, s(t)w, z(t)) dt + (s(T)w, z_T) - (s(0)w, y_0) \\ & = \int_0^T (s(t)w, e(t)) dt. \end{aligned}$$

Insbesondere gilt für alle $\varphi \in \mathcal{D}([0, T])$

$$(9.11) \quad - \int_0^T (w, \dot{\varphi}(t)z(t))dt + \int_0^T (w, \varphi(t)A(t)z(t))dt = \int_0^T (w, \varphi(t)e(t))dt \quad \forall w \in V.$$

Da man in (9.11) die Elemente w vor das Integral ziehen darf (vergl. YOSHIDA, Kap. V, Abschnitt 5), folgt also

$$(9.12) \quad - \int_0^T \dot{\varphi}(t)z(t)dt + \int_0^T \varphi(t)A(t)z(t)dt = \int_0^T \varphi(t)e(t)dt \quad \forall \varphi \in \mathcal{D}([0, T]).$$

Nach Definition der Ableitung vektorwertiger Distributionen haben wir daher

$$(9.13) \quad \frac{dz}{dt} + A(t)z = e.$$

Wir wählen in (9.10) speziell solche $s \in H^1([0, T])$, für die $s(T) = 0$ gilt. Aus (9.13) folgt $\frac{dz}{dt} \in L^2(0, T; V^*)$, wir können in (9.10) also partiell integrieren. (vergl. (4.10)). Unter Ausnutzung von (9.13) ergibt sich dann aus (9.10):

$$(9.14) \quad (s(0)w, z(0) - y_0) = 0 \quad \forall w \in V.$$

Als nächstes wählen diejenigen s , für die $s(0) = 0$ gilt und können ebenfalls partiell integrieren. Dann folgt

$$(9.15) \quad (s(T)w, z_T - z(T)) = 0 \quad \forall w \in V.$$

Aus (9.14) und (9.15) schließen wir

$$(9.16) \quad z(0) = y_0, \quad z(T) = z_T.$$

Damit ist Aussage a) bewiesen.

Sei $e_k = e$ für alle $k \in \mathbb{N}$. Dann setzen wir $y_{mn} := y_{mn}^k$ und nach a) konvergiert also $\{y_{m_i n_i}\}$ schwach gegen z und $\{y_{m_i n_i}(T)\}$ schwach gegen $z(T)$. Da die Lösung der Differentialgleichung (1.37), (1.38) eindeutig bestimmt ist, konvergiert sogar $\{y_{mn}\}$ schwach gegen z und $\{y_{mn}(T)\}$ schwach gegen $z(T)$. Wir wollen zeigen, daß man hieraus auf Normkonvergenz schließen kann. Mit (9.4) ergibt sich

$$(9.17) \quad \int_0^T a(t, y_{mn}(t) - z(t), y_{mn}(t) - z(t))dt + \frac{1}{2} \|y_{mn}(T) - z(T)\|^2 + \frac{1}{2} \|y_{mn}(0) - y_0\|^2$$

$$\begin{aligned}
 &= \int_0^T a(t, y_{mn}(t), y_{mn}(t)) dt + \frac{1}{2} |y_{mn}(T)|^2 + \frac{1}{2} |y_{mn}(0)|^2 - (y_{mn}(0), y_0) \\
 &- \int_0^T a(t, y_{mn}(t), z(t)) dt - \int_0^T a(t, z(t), y_{mn}(t)) dt + \int_0^T a(t, z(t), z(t)) dt \\
 &- (y_{mn}(T), z(T)) + \frac{1}{2} |z(T)|^2 + \frac{1}{2} |y_0|^2 \\
 &= \int_0^T (y_{mn}(t), e(t)) dt - \int_0^T a(t, y_{mn}(t), z(t)) dt - \int_0^T a(t, z(t), y_{mn}(t)) dt \\
 &+ \int_0^T a(t, z(t), z(t)) dt - (y_{mn}(T), z(T)) + \frac{1}{2} |z(T)|^2 + \frac{1}{2} |z(0)|^2.
 \end{aligned}$$

Also konvergiert der Term (9.17) gegen

$$\int_0^T (z(t), e(t)) dt - \int_0^T a(t, z(t), z(t)) dt - \frac{1}{2} |z(T)|^2 + \frac{1}{2} |z(0)|^2.$$

Wegen (9.13) und (9.16) verschwindet dieser Term. Andererseits majorisiert (9.17) die Summe

$$\propto \int_0^T \|y_{mn}(t) - z(t)\|^2 dt + \frac{1}{2} |y_{mn}(0) - y_0|^2,$$

die also auch gegen Null strebt. Damit ist die Behauptung bewiesen.]

Bemerkung: Es ist klar, daß der oben gegebene Beweis gleichzeitig die Existenz einer Lösung der parabolischen Differentialgleichung mitliefert.

Wir benötigen ein Satz 9.1 b) entsprechendes Ergebnis für die adjungierte Differentialgleichung. Für $e^* \in L^2(0, T; V^*)$ betrachten wir die Gleichung

$$(9.18) \quad -\frac{dw^*}{dt} + A^*(t)w^* = e^*,$$

$$(9.19) \quad w^*(T) = 0.$$

Wir konstruieren Näherungslösungen $w_{mn}^* \in W_{mn}$ vermöge

$$(9.20) \quad -\left[z, \frac{dw_{mn}^*}{dt}\right] + \int_0^T a(t, w_{mn}^*(t), z(t)) dt + (z(T), w_{mn}^*(T)) = [z, e^*] \quad \forall z \in W_{mn}.$$

Satz 9.2. Die Gleichung (9.20) ist eindeutig durch ein $w_{mn}^* \in W_{mn}$ erfüllbar. Die Folge der w_{mn}^* konvergiert in der Norm von $L^2(0, T; V)$ gegen die Lösung w^* der Differentialgleichung (9.18), (9.19). Die Folge $w_{mn}^*(T)$ konvergiert in der Normtopologie von H gegen 0.

Beweis: Wir setzen

$$\tilde{w}_{mn}^*(t) := w_{mn}^*(T-t), \quad \tilde{z}(t) := z(T-t).$$

Durch den Übergang von t nach $T-t$ erhalten wir aus S_n, W_{mn} Teilräume $\tilde{S}_n, \tilde{W}_{mn}$, für die ebenfalls die Bedingungen (9.2), (9.3) erfüllt sind. (9.20) ist äquivalent mit

$$(9.21) \quad \left[\tilde{z}, \frac{d\tilde{w}_{mn}^*}{dt}\right] + \int_0^T a(t, \tilde{w}_{mn}^*(t), \tilde{z}(t)) dt + (\tilde{z}(0), \tilde{w}_{mn}^*(0)) = [\tilde{z}, e^*] \quad \forall \tilde{z} \in \tilde{W}_{mn}.$$

Die $\tilde{w}_{mn}^*$ konstruieren sich also aus dem Galerkin-Verfahren für die Differentialgleichung

$$\frac{d\tilde{w}^*}{dt} + A(t)\tilde{w}^* = e^*, \quad \tilde{w}^*(0) = 0.$$

Beachtet man, daß $w^*(t) := \tilde{w}^*(T-t)$ die Gleichung (9.18), (9.19) löst, so folgt mit Satz 9.1. die Behauptung. $\square$

Zum Schluß deuten wir noch an, wie man in einem konkreten Beispiel die Teilräume V_m wählen kann.

Sei $V = H^1(\Omega), \Omega \subset \mathbb{R}^2$. Zu jedem $h \in (0, 1)$ gibt es einen linearen Raum $S_h \subset H^2(\Omega)$ von finiten Elementen, der die Eigenschaft hat, daß zu jedem $v \in H^2(\Omega)$ ein $\phi \in S_h$ mit

$$\|v - \phi\|_{H^1(\Omega)} \leq C h \|v\|_{H^2(\Omega)}$$

existiert. Dabei ist C eine von v und h unabhängige Konstante. Wähle eine Nullfolge $\{h_m\} \subset (0, 1)$ und setze $V_m = S_{h_m}$. Da $H^2(\Omega)$ dicht in $H^1(\Omega)$ ist, folgt mit einer ähnlichen Aussage wie der von Lemma 8.1 die Voraussetzung (9.2). Für die Teilräume S_n kann man analog vorgehen.

§10. Konvergenzaussagen für das parabolische Kontrollproblem

10.1. Konvergenz der Extremalwerte

Wir sind nun in der Lage, die Ergebnisse aus §6 auf die Diskretisierungen des parabolischen Kontrollproblems anzuwenden.

Satz 10.1. Existiere zu f^* aus (4.5) eine konvexe Funktion $f_1 \in C(B^*)$ mit $f_1(v^*) = F(v^*)$ für alle $v^* \in \text{dom } f^*$. Falls Voraussetzung (2.2) für alle U_{ad}^1 erfüllt ist (U_{ad}^1 anstelle von U_{ad}), gilt unter den Voraussetzungen (8.3), (8.4), (9.2) und (9.3) für die in §7 eingeführten diskreten Kontrollprobleme (P_{mn}^1)

$$(10.1) \quad \min(P) = \lim_{m,n,l \rightarrow \infty} \min(P_{mn}^1).$$

Beweis: Wir haben die Voraussetzungen von Satz 6.1 nachzuprüfen. Die Stetigkeit der Zielfunktionen folgt sofort aus unseren Voraussetzungen und der Tatsache, daß eine überall endliche konkave Funktion stetig ist, wenn sie in einer Nullumgebung nach unten beschränkt ist. Die Gültigkeit von (V0) ergibt sich aus Satz 7.2.

Zum Nachweis von (V1) wählen wir ein für (P) zulässiges (y, u) . Es gibt eine Folge $\{u_l\}$ mit $u_l \in U_{ad}^1$ für alle $l \in \mathbb{N}$ und $u = \lim_{l \rightarrow \infty} u_l$. Wir haben in §7 gezeigt, daß für beliebiges $g \in L^2(0, T; V^*)$ und $\eta \in H$ die Gleichung

$$b(z, y) = (z(0), \eta) + [z, g] \quad \forall z \in W_{mn}$$

durch genau ein $y \in W_{mn}^2$ lösbar ist. Also können wir durch $\Lambda_{mn}(g, \eta) := y$ einen linearen Operator von $L^2(0, T; V^*) \oplus H$ nach $L^2(0, T; V)$ definieren. Λ_{mn} ist stetig, denn wegen (9.5) gilt

$$\alpha \|\Lambda_{mn}(g, \eta)\|_{L^2(0, T; V)}^2 \leq \frac{1}{\alpha} \|g\|_{L^2(0, T; V^*)}^2 + |\eta|^2.$$

Entsprechend definieren wir zu rechter Seite g und Anfangswert η einen Operator Λ durch $\Lambda(g, \eta) := y$, wobei y die Lösung der Differentialgleichung (1.37), (1.38) sei. Dann folgt aus Satz 9.1

$$(10.2) \quad \Lambda(g, \eta) = \lim_{m,n \rightarrow \infty} \Lambda_{mn}(g, \eta).$$

Nach dem Prinzip der gleichmäßigen Beschränktheit gibt es daher eine Konstante $c > 0$,

so daß für alle $m, n \in \mathbb{N}$

$$(10.3) \quad \|\bigwedge_{mn}\| \leq c$$

gilt. Wir setzen $y_{mn}^1 := \bigwedge_{mn}(e+Bu_1, y_0)$. Dann folgt (mit $\|\cdot\| := \|\cdot\|_{L^2(0,T;V)}$)

$$(10.4) \quad \|y_{mn}^1 - y\| \leq \|\bigwedge_{mn}(e+Bu_1, y_0) - \bigwedge_{mn}(e+Bu, y_0)\| \\ + \|\bigwedge_{mn}(e+Bu, y_0) - \bigwedge(e+Bu, y_0)\|.$$

Der erste Summand wird durch

$$\|\bigwedge_{mn}\| \|B\| \|u_1 - u\|_U$$

majorisiert. Wegen (10.2) und (10.3) konvergiert also $\{y_{mn}^1\}$ gegen y .

(V1_{*}) folgt sofort wegen Satz (9.2). (V2) und (V3) sind hier trivial. Da es eine Konstante c_1 mit

$$(10.5) \quad \langle u_1, Nu_1 + B^* w^* \rangle_U \leq c_1 + \langle u, Nu + B^* w^* \rangle_U \quad \forall 1 \in \mathbb{N}$$

gibt, haben wir auch (V3_{*}). Setze

$$g(u) := \langle u, Nu + B^* w^* \rangle_U.$$

Für alle u, v mit $\|u\|_U \leq m, \|v\|_U \leq m, m > 0$ gilt

$$|g(u) - g(v)| = |\langle u - v, N(u+v) + B^* w^* \rangle_U| \\ \leq \|u - v\|_U [\|N(u+v)\|_U + \|B^* w^*\|_U] \\ \leq \|u - v\|_U [2m\|N\| + \|B^* w^*\|_U].$$

g ist also auf beschränkten Mengen gleichmäßig stetig. Berücksichtigen wir noch (10.5), so können wir auf g und die Mengen U_{ad}^1, U_{ad} Lemma 6.3 anwenden. Damit folgt (V2_{*}) und Satz 10.1 ist bewiesen. $\square$

Bemerkung: Die f' betreffende Voraussetzung in Satz 10.1 ist in den Beispielen aus §4.3 erfüllt.

10.2. Konvergenz der Optimallösungen

Für den Rest des Paragraphen bezeichnen wir mit (y,u) eine Optimallösung von (P) und mit (y_{mn}, u_1) eine Optimallösung von (P_{mn}^1) .

Satz 10.2. Sei f nach unten beschränkt, wenn U_{ad} unbeschränkt ist. Falls eine Folge $\{u_i\}$ mit $u_i \in U_{ad}^1$ für alle $i \in \mathbb{N}$ in der schwachen Topologie gegen ein gewisses $u \in U$ konvergiert, gelte $u \in U_{ad}$. Hierbei sei $\{1_i\}$ eine Teilfolge von $\{i\}_{i \in \mathbb{N}}$. Zusätzlich zu diesen Annahmen seien die Voraussetzungen von Satz 10.1 erfüllt. Dann konvergiert eine Teilfolge von $\{(y_{mn}, u_1)\}$ schwach gegen eine Optimallösung (y,u) von (P) .

Beweis: Falls N koerzitiv ist, gilt

$$f(Cy_{mn} - z_d) + \langle u_1, Nu_1 \rangle_U \leq f(Cy_{mn} - z_d) + \nu \|u_1\|_U^2.$$

Nach Satz 10.1 gibt es eine Konstante c mit

$$c + \min(P) \geq \min(P_{mn}^1)$$

für alle $m, n, l \in \mathbb{N}$. Unter Berücksichtigung von Voraussetzung (2.2) sieht man also, daß die Folge $\{u_1\}$ in einer beschränkten Menge liegt. Aus Satz 10.1 b) und Satz 6.2 folgt die Behauptung, wenn man beachtet, daß eine stetige konvexe Funktion stets schwach nach unten halbstetig ist. $\square$

Man kann die Konvergenzaussage von Satz 10.2 verschärfen, wenn man weiß, daß die Zielfunktion strikt konvex ist.

Satz 10.3. Sei N koerzitiv, und seien die Voraussetzungen von Satz 10.2 erfüllt. Dann sind die Optimallösungen (y_{mn}, u_1) und (y,u) eindeutig bestimmt, und $\{(y_{mn}, u_1)\}$ konvergiert schwach gegen (y,u) .

Im Fall der quadratischen Probleme ist es im koerzitiven Fall sogar möglich, auf Normkonvergenz zu schließen.

Satz 10.4. Seien $\mathcal{B}$ und f wie in Beispiel 1 von §4.3. N sei koerzitiv und die Voraussetzungen von Satz 10.2 seien erfüllt. Dann konvergiert $\{(y_{mn}, u_1)\}$ in der Normtopologie gegen (y,u) .

Beweis: Aus Satz 10.1 folgt

$$\|C y_{mn} - z_d\|_{\mathcal{H}}^2 + \langle u_1, N u_1 \rangle_U \longrightarrow \|C y - z_d\|_{\mathcal{H}}^2 + \langle u, N u \rangle_U.$$

Also haben wir auch

$$\begin{aligned} \|C y_{mn}\|_{\mathcal{H}}^2 - 2 \langle C y_{mn}, z_d \rangle_{\mathcal{H}} + \|z_d\|_{\mathcal{H}}^2 + \langle u_1, N u_1 \rangle_U \\ \longrightarrow \|C y\|_{\mathcal{H}}^2 - 2 \langle C y, z_d \rangle_{\mathcal{H}} + \|z_d\|_{\mathcal{H}}^2 + \langle u, N u \rangle_U. \end{aligned}$$

Mit Satz 10.3 folgt hieraus

$$(10.6) \quad \|C y_{mn}\|_{\mathcal{H}}^2 + \langle u_1, N u_1 \rangle_U \longrightarrow \|C y\|_{\mathcal{H}}^2 + \langle u, N u \rangle_U.$$

Der Term

$$\begin{aligned} \|C(y - y_{mn})\|_{\mathcal{H}}^2 + \langle u - u_1, N(u - u_1) \rangle_U \\ = \|C y\|_{\mathcal{H}}^2 + \langle u, N u \rangle_U + \|C y_{mn}\|_{\mathcal{H}}^2 + \langle u_1, N u_1 \rangle_U - 2 \langle C y_{mn}, C y \rangle_{\mathcal{H}} - 2 \langle u_1, N u \rangle_U \end{aligned}$$

konvergiert also gegen Null, majorisiert andererseits $\|u - u_1\|_U^2$. Also konvergiert $\{u_1\}$ stark gegen u . Die Normkonvergenz der y_{mn} folgt wie in Beweis von Satz 10.1 aus (10.3). $\square$

Satz 10.5. Sei $N = 0$ und f wie in Beispiel 1 von §4.3. Unter den Voraussetzungen von Satz 10.2 konvergiert dann $\{C y_{mn}\}$ in der Normtopologie gegen $C y$.

Beweis: Seien (y, u) und $(\tilde{y}, \tilde{u})$ Optimallösungen von (P). Da $\mathcal{H}$ als Hilbert-Raum strikt normiert ist, folgt aus $\|C y - z_d\|_{\mathcal{H}}^2 = \|C \tilde{y} - z_d\|_{\mathcal{H}}^2$ sofort die Gültigkeit von

$$(10.7) \quad C y = C \tilde{y}.$$

Wie im Beweis von Satz 10.2 folgt daher, daß $\{C y_{mn}\}$ schwach gegen $C y$ konvergiert. Wegen Satz 10.1 konvergiert also auch $\{\|C y_{mn}\|_{\mathcal{H}}^2\}$ gegen $\|C y\|_{\mathcal{H}}^2$, woraus sich die Behauptung unmittelbar ergibt. $\square$

Ist N nicht koerzitiv, so kann man nur in speziellen Fällen auf Normkonvergenz schließen. Sei $U := L^2(Q)$ ($Q :=]0, T[\times \Omega$) und

$$(10.8) \quad U_{ad} := \{u \in L^2(Q) \mid |u| \leq m\}.$$

Satz 10.6. Sei U_{ad} gemäß (10.8) gegeben und sei $U_{ad}^l \subset U_{ad}$ für alle $l \in \mathbb{N}$. Für jede optimale Steuerung u von (P) gelte $|u| = m$. Dann ist u eindeutig bestimmt, und unter den Voraussetzungen von Satz 10.1 konvergiert $\{u_l\}$ in der Normtopologie gegen u .

Beweis: Es ist klar, daß u eindeutig bestimmt ist und aus Satz 10.2 folgt, daß $\{u_l\}$ schwach gegen u konvergiert. Die Normkonvergenz von $\{u_l\}$ folgt aus LEVITIN-POLJAK [24], Theorem 4. $\square$

§11. Ein numerisches Beispiel

11.1. Das diskrete Problem in Matrizenschreibweise

In diesem Paragraphen deuten wir an, wie man das diskretisierte Kontrollproblem (P_{mn}^1) aus §7.1 numerisch lösen kann und teilen die numerischen Ergebnisse mit, die sich bei der praktischen Durchrechnung des Transmissions-Kontrollproblems aus §2.2 ergeben haben.

Zunächst ist es erforderlich, im Problem (P_{mn}^1) für die Räume W_{mn} , U_1 Basen einzuführen, um auf diese Weise zu einem auf dem $\mathbb{R}^k$ definierten Optimierungsproblem zu gelangen (mit geeignetem k). Sei

$$(11.1) \quad \begin{aligned} U_1 &= \text{lin}\{u_1, \dots, u_l\}, \\ S_n &= \text{lin}\{s_1, \dots, s_n\}, \quad V_m = \text{lin}\{v_1, \dots, v_m\}. \end{aligned}$$

Hierbei nehmen wir an, daß $l = \dim U_1$, $n = \dim S_n$ und $m = \dim V_m$ gilt.

Für $i, \mu = 1, \dots, m; j, \nu = 1, \dots, n$ definieren wir

$$(11.2) \quad m_{ij\mu\nu} := \int_0^T (s_j(t)w_i, \dot{s}_\nu(t)w_\mu) dt + \int_0^T a(t, s_j(t)w_i, s_\nu(t)w_\mu) dt + (s_j(0)w_i, s_\nu(0)w_\mu).$$

Hieraus erhalten wir $m \times n$ -Matrizen

$$(11.3) \quad m_{i\mu} := (m_{ij\mu\nu})$$

und eine $(mn) \times (mn)$ -Blockmatrix

$$(11.4) \quad M := (m_{i\mu}).$$

Aus Lemma 7.1 folgt, daß M nichtsingulär ist. Falls die Bilinearformen $a(t; \cdot, \cdot)$ von t unabhängig sind, gilt

$$(11.5) \quad m_{ij\mu\nu} = (w_i, w_\mu) \int_0^T (s_j(t), \dot{s}_\nu(t)) dt + a(w_i, w_\mu) \int_0^T s_j(t)s_\nu(t) dt + s_j(0)s_\nu(0)(w_i, w_\mu).$$

Für $i=1, \dots, m; j=1, \dots, n$ und $\lambda = 1, \dots, l$ sei

$$(11.6) \quad \gamma_{0ij} := (s_j(0)w_i, y_0), \quad \bar{e}_{ij} = \int_0^T (s_j(t)w_i, e(t)) dt,$$

$$(11.7) \quad d_{ij\lambda} := \int_0^T (s_j(t)w_i, Bv_\lambda(t)) dt.$$

Hieraus ergeben sich in naheliegender Weise Vektoren η_0 , $\bar{e} \in \mathbb{R}^{mn}$ und eine $(mn) \times 1$ -Matrix D , die sich aus Blockmatrizen

$$(11.8) \quad d_i := (d_{ij\lambda})$$

zusammensetzt. Stellen wir $y \in W_{mn}$ und $u \in U_1$ in der Form

$$(11.9) \quad y = \sum_{i=1}^m \sum_{j=1}^n \eta_{ij} s_j w_i, \quad u = \sum_{\lambda=1}^1 \xi_\lambda u_\lambda,$$

dar und definieren wir Vektoren $\eta := \eta_{ij} \in \mathbb{R}^{mn}$, $\xi := (\xi_\lambda) \in \mathbb{R}^1$, so ist (7.8) also äquivalent mit

$$(11.10) \quad M\eta = \bar{e} - \eta_0 + D\xi.$$

Da die Zielfunktion des Transmissions-Kontrollproblems quadratisch ist, wollen wir uns auch hier auf den quadratischen Fall beschränken. Seien also f und B wie in Beispiel 1 von §4.3. Wir definieren

$$(11.11) \quad g_{i\mu} := (\langle C(s_j w_i), C(s_\nu w_\mu) \rangle_{\mathcal{X}}), \quad g := (\langle C(s_j w_i), z_d \rangle_{\mathcal{X}})$$

und eine Blockmatrix

$$(11.12) \quad G := (g_{i\mu}).$$

Man rechnet leicht nach, daß

$$\|Cy - z_d\|_{\mathcal{X}}^2 = \langle \eta, G\eta \rangle - 2\langle \eta, g \rangle + \|z_d\|_{\mathcal{X}}^2$$

gilt. Hierbei bezeichne unabhängig von $k \in \mathbb{N}$ $\langle \cdot, \cdot \rangle$ das gewöhnliche euklidische Skalarprodukt des $\mathbb{R}^k$. Schließlich setzen wir

$$(11.13) \quad H := (\langle u_\lambda, Nu_\lambda \rangle),$$

$$(11.14) \quad \gamma := \bar{e} - \eta_0, \quad z_0 := \|z_d\|_{\mathcal{X}}^2.$$

Dann ist also (P_{mn}^1) mit folgendem quadratischen Programmierungsproblem äquivalent.

$$(P_{mn}^1) \text{ Minimiere } \langle \eta, G\eta - 2g \rangle + \langle \xi, H\xi \rangle + z_0$$

über alle $\eta \in \mathbb{R}^{mn}$ und $\xi \in \bar{U}_{ad}^1$ unter der Nebenbedingung

$$(11.15) \quad M\eta = \gamma + D\xi.$$

Wie wir in den Beispielen von §8.2 gesehen haben, wird $\tilde{U}_{ad}^1$ in vielen Fällen durch eine endliche Anzahl linearer Ungleichungen beschrieben.

11.2. Die numerischen Ergebnisse

Man kann in dem quadratischen Programmierungsproblem ($\tilde{P}_{mn}^1$) durch Inversion der Matrix M die Variablen eliminieren. In dem verbleibenden Problem ist dann, falls eine geeignete Struktur von $\tilde{U}_{ad}^1$ vorausgesetzt wird, eine quadratische Zielfunktion unter gewissen Ungleichungsnebenbedingungen zu minimieren. Dieses Problem kann mit einem geeigneten Minimierungsverfahren behandelt werden, so z.B. mit der Methode des projizierten Gradienten, vergleiche etwa WOLFE [39] oder CEA [6]. Mit Hilfe einer derartigen Vorgehensweise sind die folgenden numerischen Ergebnisse gewonnen worden. Zur Matrixinversion diente dabei das Jordan-Verfahren.

Wir betrachten erneut das Transmissions-Kontrollproblem aus §2.2. Dabei sei $\Omega_1 :=]a, b[$ und $\Omega_2 :=]b, c[$. Für die in Ω_1 und Ω_2 gegebenen Konstanten wählen wir $k_1 = \lambda_1 = 0.004$ und $k_2 = \lambda_2 = 0.12$. Weiter sei $a=0$, $b=3$, $c=8$, $T=50$, $y_{01}=y_{02}=0$, $m=1$ und $z_d=25$. Um die Basisfunktionen der approximierenden Teilräume definieren zu können, führen wir Zerlegungen

$$(11.16) \quad a = x_0 < x_1 < \dots < x_m = b < \dots < x_{m_0} = b,$$

$$(11.17) \quad 0 = t_0 < t_1 < \dots < t_{n_0} = T,$$

$$(11.18) \quad b = x'_0 < x'_1 < \dots < x'_1 = c, \quad 0 = t'_0 < t'_1 < \dots < t'_1 = T$$

ein. W_m bzw. S_n sei dann der zur Zerlegung (11.16) bzw. (11.17) gehörige Raum von kubischen Splinefunktionen. Basen dieser Räume wählen wir gemäß WERNER-SCHABACK [38], Kap. III, §3. Die Ansatzfunktionen für die Steuerungen werden wie in Beispiel 1 von §8.2 konstruiert.

Alle Rechnungen über die wir im folgenden berichten, wurden auf einer Telefunken TR 440-Anlage durchgeführt. Die Zerlegungen (11.16), (11.17) blieben bei allen Rechnungen fest. Wir haben $m_0=12$ und $n_0=9$ gesetzt und für die Stützstellen aus (11.16) bzw. (11.17) die folgenden Werte gewählt:

0, 1, 2, 2.4, 2.7, 2.9, 3, 3.5, 4, 5, 6, 7, 8;

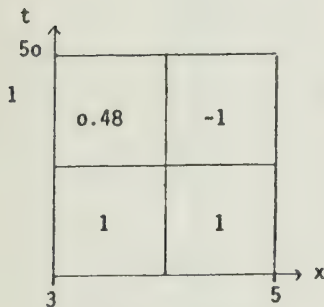
0, 10, 20, 25, 30, 33, 38, 40, 45, 50.

Die Werte für die Zerlegungen (11.18) ergeben sich aus Tabelle 1.

Beispiel Nr.	x'_i	t'_i
1	3, 5.5, 8	0, 25, 50
2	3, 4.66, 6.33, 8	0, 16.66, 33.33, 50
3	3, 4.25, 5.5, 6.75, 8	0, 12.5, 25, 37.5, 50
4	3, 4, 5, 6, 7, 8	0, 25, 33, 41, 50

Tab. 1

Die hieraus errechneten diskreten Optimallösungen und Optimalwerte sind in Tabelle 2 wiedergegeben. Schließlich ist in Abbildung 2 der Verlauf von Cy_{mn}^1 graphisch dargestellt.



2	-0.23	1	0.19
	1	1	-0.30
	1	1	1

3	0.42	0.19	-0.01	0.00
	1	-0.93	-0.32	-0.03
	1	1	1	1
	1	1	1	1

4	0.11	0	0	0	0
	-0.02	0.27	0.08	0.01	0
	1	0.50	-0.09	0.04	0.02
	1	1	1	1	1

Beispiel Nr.	$\inf (P_{mn}^1)$
1	6709.74
2	6667.22
3	6664.61
4	6660.21

Tab. 2

Wie man aus Tabelle 2 und Abbildung 2 erkennt, bestätigen sich die Aussagen von Satz 10.1 und Satz 10.5 auch numerisch. Abbildung 2 legt die Vermutung nahe, daß für die exakte Optimallösung y die Funktion Cy erst linear ansteigt bis sie den Wert z_d erreicht hat und dann konstant mit dem Wert z_d weiter verläuft. Diese Annahme ist auch physikalisch sinnvoll.

Weiter zeigt es sich, daß das Bang-Bang-Prinzip für die optimale Steuerung u

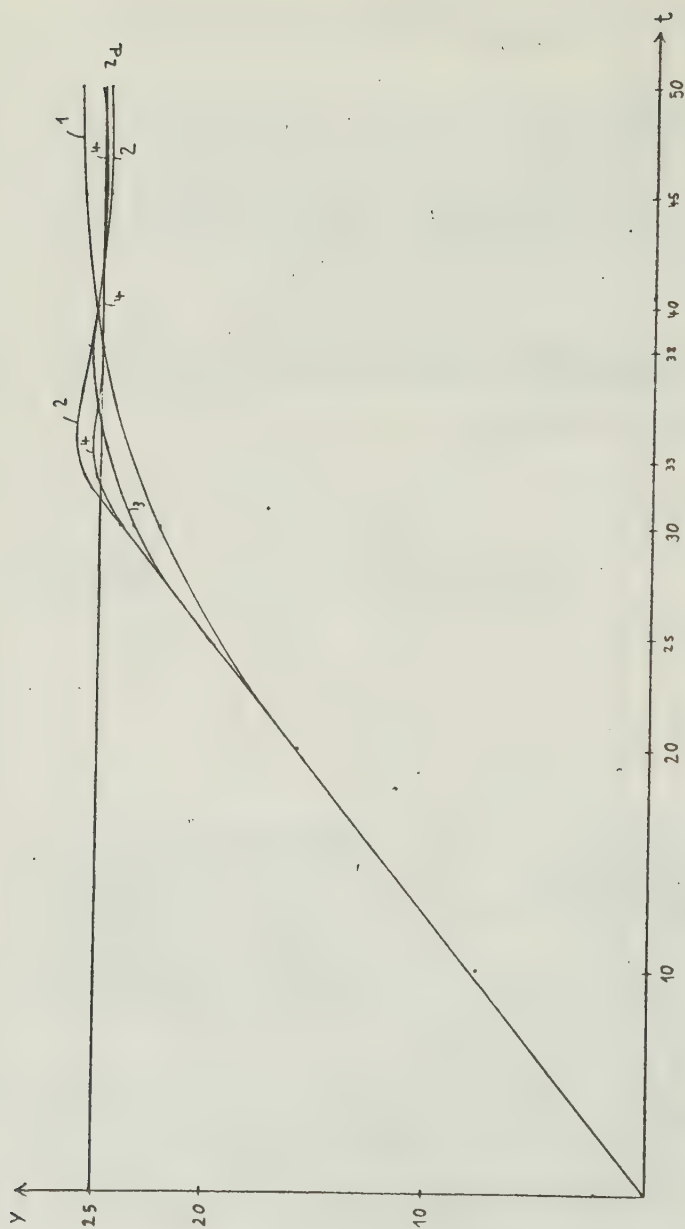


Abb. 2

offenbar nicht erfüllt ist, denn einige Werte von u pendeln sich in der Nähe von Null ein und der Wert -1 wird, abgesehen vom 1. Beispiel, nicht mehr angenommen. Demnach wäre Satz 10.6 nicht anwendbar. Tabelle 2 gibt keinen deutlichen Hinweis, ob die Folge der u_j in der Norm von $L^2(Q_2)$ konvergiert. Schließlich bemerken wir, daß die mit dem Gradientenverfahren berechnete Iterationsfolge nur sehr langsam gegen die Näherungslösung u_j konvergierte. Vermutlich liegt dies an der fehlenden Definitheit der kontinuierlichen Zielfunktion.

LITERATURVERZEICHNIS

- [1] ADAMS, R.A.: Sobolev-Spaces, Academic Press, New York - San Francisco - London, 1975.
- [2] BABUSKA, I., AZIZ, A.K.: The Mathematical Foundations of the Finite Elements Method with Applications to Partial Differential Equations, Editor: A.K. Aziz, Academic Press, New York - London, 1972.
- [3] BENSOUSSAN, A., BOSSAVIT, A., NEDELEC, J.C.: Approximation des problèmes de contrôle optimal, Cahiers de l'IRIA, no. 2, 107-176 (1970).
- [4] BOSARGE, W.E., JOHNSON, O.G., SMITH, C.L.: A Ritz-Galerkin spline approximation to the linear parabolic regulator control problem, Preprints of Papers for IFAC-Symposium on the Control of Distributed Parameter Systems, Banff, Canada, Juni 1971.
- [5] BUTKOVSKIY, A.G.: Distributed control systems, Elsevier, New York, 1969.
- [6] CEA, J.: Optimisation, Théorie et Algorithmes, Dunod, Paris, 1971.
- [7] COLLATZ, L.: Funktionalanalysis und numerische Mathematik, Springer-Verlag, Berlin - Heidelberg - New York, 1968.
- [8] DANIEL, J.W.: The Approximate Minimization of Functionals, Prentice - Hall, Inc., Englewood Cliffs, 1971.
- [9] DOUGLAS, J., DUPONT, T.: Galerkin methods for parabolic equations, SIAM J. Numer. Anal., Vol 7., Nr. 4, Dez. 1970.
- [10] ECKHARDT, U.: Quadratic programming by successive overrelaxation, Berichte der Kernforschungsanlage Jülich - Nr. 1064, April 1974.
- [11] GLASHOFF, K., GUSTAFSON, S.-A.: Numerical treatment of a parabolic boundary - value control problem, The Royal Institute of Technology, Stockholm, Sweden, Report TRITA-NA-7409, 1974.
- [12] GLASHOFF, K., KRABS, W.: Konvergenz der Linienmethode bei einem parabolischen Rand-Kontrollproblem, Technische Hochschule Darmstadt, Preprint Nr. 88, 1973.
- [13] GLASHOFF, K., KRABS, W.: Dualität und Bang-Bang-Prinzip bei einem parabolischen Rand-Kontrollproblem, Technische Hochschule Darmstadt, Preprint Nr. 155, August 1974.

- [14] GLASHOFF, K., WECK, N.: Boundary control of parabolic differential equations in arbitrary dimensions: supremum-norm problems, Technische Hochschule Darmstadt, Preprint Nr. 199, April 1975.
- [15] GRUVER, W.A.: Modeling of heat conduction in composite solids, Vortragsausarbeitung, Technische Hochschule Darmstadt, 1973.
- [16] HIRZEBRUCH, F., SCHARLAU, W.: Einführung in die Funktionalanalysis, Bibliographisches Institut, Mannheim, 1970.
- [17] HORVATH, J.: Topological Vektor Spaces and Distributions, Addison Wesley Publishing Company, Inc., Massachusetts - Palo Alto - London, 1966.
- [18] HUET, D.: Distributions and Sobolev Spaces. An Elementary Introduction and a Few Applications, University of Maryland, Lecture Note Nr. 6, 1970.
- [19] KRABS, W.: Zur stetigen Abhängigkeit des Extremalwertes eines konvexen Optimierungsproblems von einer stetigen Änderung des Problems, ZAMM 52, 359-368, (1972).
- [20] KRABS, W.: Stetigkeitsfragen bei der Diskretisierung konvexer Optimierungsprobleme, Sonderdruck aus ISNM Vol. 17, 1973.
- [21] KRABS, W.: Stabilität und Stetigkeit bei nichtlinearer Optimierung, in: Methoden des Operations Research XVII, Herausgeber: R. Henn, 1973.
- [22] KRABS, W.: Zur Berechnung des Extremalwertes bei einem parabolischen Randkontrollproblem, Technische Hochschule Darmstadt, Preprint Nr. 141, Mai 1974.
- [23] KRABS, W., WECK, N.: Über ein Kontrollproblem in der Wärmeleitung, Technische Hochschule Darmstadt, Preprint Nr. 108, Januar 1974.
- [24] LEVITIN, E.S., POLJAK, B.T.: Convergence of minimizing sequences in conditional extremum problems, Soviet Math. Dokl., Tom 168, No. 5, 764-767, (1966).
- [25] LIONS, J.L.: Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin - Heidelberg - New York, 1971.
- [26] LIONS, J.L.: Some Aspects of the Optimal Control of Distributed Parameter Systems, Regional Conference Series in Applied Mathematics No. 6, SIAM, Philadelphia - Pennsylvania, 1972.

- [27] LIONS, J.L.: Various topics in the theory of optimal control of distributed parameter systems, in: Lecture Notes in Economics and Mathematical Systems 105, Part I, Herausgeber: Kirby, B.J., Springer-Verlag, Berlin-Heidelberg, New-York 1974.
- [28] LIONS, J.L.: Problèmes aux Limites dans les Equations aux Dérivées Partielles, Les Presses de l'Université de Montréal, 1965.
- [29] LIONS, J.L.: Equations Différentielles Opérationnelles, Springer-Verlag, Berlin - Göttingen - Heidelberg, 1961.
- [30] LIONS, J.L.: Contribution à un problème de M. Picone, Ann. Mat. Pura Appl. 4, XLI, 201-219, (1956).
- [31],[32] LIONS, J.L., MAGENES, E.: Non-Homogeneous Boundary Value Problems and Applications, Vol. I, II, Springer-Verlag, Berlin- Heidelberg- New-York, 1972.
- [33] MOREAU, J.-J.: Fonctionelles Convexes, vervielfältigte Vorlesungsausarbeitung, Séminaire sur les Equations aux Dérivées Partielles, Collège de France, 1967.
- [34] ROCKAFELLAR, R.T.: Duality and stability in extremum problems involving convex functions, Pacific J. Math. 21, 167-187, (1967).
- [35] ROCKAFELLAR, R.T.: Convex Analysis, Princeton University Press, Princeton 1969.
- [36],[37] SCHWARTZ, L.: Théorie des Distributions, Vol. I, II, Hermann, Paris 1957.
- [38] WERNER, H., SCHABACK, R.: Praktische Mathematik II, Springer-Verlag, Berlin - Heidelberg - New-York, 1972.
- [39] WOLFE, P.: Methods of Nonlinear Programming, in: Nonlinear Programming, Herausgeber: Abadie, J., Interscience, John Wiley, New-York 1967.
- [40] YOSHIDA, K.: Functional Analysis, 4. Auflage, Springer-Verlag, Berlin - Heidelberg - New-York, 1974.

LEBENS LAUF

Am 30. September 1947 wurde ich in Hamburg geboren als Sohn des Chemotechnikers Günther Mackenroth und der Laborantin Liesel Mackenroth, geb. Boldt.

1974 heiratete ich die Studentin Gabriela Friedrichson.

Von 1954 bis 1964 besuchte ich die Volks - und Mittelschule Meerweinstraße 26. Danach wechselte ich zum Gymnasium an der Wartenau über, wo ich im Januar 1967 das Abitur bestand. Anschließend leistete ich einen achtzehnmonatigen Grundwehrdienst beim Bundesgrenzschutz ab.

Im Wintersemester 1968/69 begann ich das Studium der Mathematik an der Universität Hamburg. 1970 wurde ich in die Studienstiftung des Deutschen Volkes aufgenommen. Die Diplomvorprüfung legte ich im Herbst 1970 ab, und im Sommer 1973 bestand ich die Diplomhauptprüfung mit einer Arbeit über konvexe Variations - und Steuerungsprobleme.

Seit dem Juli 1973 bin ich im Fachbereich Mathematik der Johannes Gutenberg Universität in Mainz als wissenschaftlicher Mitarbeiter von Herrn Professor Dr. H.N. Mülthei tätig.

